

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
17 October 2002 (17.10.2002)

PCT

(10) International Publication Number
WO 02/081749 A2

- (51) International Patent Classification⁷: **C12Q 1/68**
- (21) International Application Number: **PCT/US01/51652**
- (22) International Filing Date: 26 October 2001 (26.10.2001)
- (25) Filing Language: **English**
- (26) Publication Language: **English**
- (30) Priority Data:
09/699,243 27 October 2000 (27.10.2000) **US**
- (71) Applicant (for all designated States except US): **UNIVERSITY OF SOUTHERN CALIFORNIA [US/US]; 3716 Hope Street #313, Los Angeles, CA 90007-4344 (US).**
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): **MARKL, Isabel [US/US]; 1005 Rashford Drive, Placentia, CA 92870 (US). JONES, Peter, A. [US/US]; 4645 Lasheart Drive, La Canada, CA 91011 (US). TOMIGAHARA, Yoshitaka [JP/JP]; 2-10-2-246, Sonehigashi-machi, Toyonaka, Osaka 561-0802 (JP). LIANG, Gangning [CN/US]; 3436 Ashbourne Place, Rowland Heights, CA 91748 (US). FU, Hualin [CN/US]; 500 Norht Atlantic Boulevard, Apt. 310, Alhanbra, CA 91801 (US). CHEN, Jonathan [—/US]; 1008 South Marguerita Avenue, Apt. 1, Alhambra, CA 91803 (US).**
- (74) Agent: **DAVISON, Barry, L.; Davis Wright Tremaine, LLP, 2600 Century Square, 1501 Fourth Avenue, Seattle, WA 98101-1688 (US).**
- (81) Designated States (national): **AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.**
- (84) Designated States (regional): **ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).**
- Published:**
— without international search report and to be republished upon receipt of that report
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

(54) Title: **METHYLATION ALTERED DNA SEQUENCES AS MARKERS ASSOCIATED WITH HUMAN CANCER**

(57) Abstract: There is disclosed (103) novel methylation-altered DNA sequences ("marker sequences") that have distinct methylation patterns in cancer, compared to normal tissue. In many instances, these marker sequences represent novel sequences not found in the GenBank data base, and none of these marker sequences have previously been characterized with respect to their methylation pattern in human cancers including, but not limited to those of bladder and prostate. These (103) sequences have utility as diagnosis, prognostic and therapeutic markers in the treatment of human cancer, and as reagents in kits for detecting methylated CpG-containing nucleic acids.

WO 02/081749 A2

METHYLATION ALTERED DNA SEQUENCES AS MARKERS ASSOCIATED WITH HUMAN CANCER

5 Cross-Reference to Related Applications

This application claims priority to U.S. Patent Application Serial No. 09/699,243, filed October 27, 2000.

Technical Field of the Invention

10 The present invention relates to novel human DNA sequences that exhibit altered methylation patterns (hypermethylation or hypomethylation) in cancer patients. These novel methylation-altered DNA sequences are useful as diagnostic, prognostic and therapeutic markers for human cancer.

15 Background of the Invention

The identification of early genetic changes in tumorigenesis is a primary focus in molecular cancer research. Characterization of the nature and pattern of cancer-associated genetic alterations will allow for early detection, diagnosis and treatment of cancer. Such genetic alterations in vertebrates fall generally into one of three categories: gain or loss of
20 genetic material; mutation of genetic material; or methylation at cytosine residues in CpG dinucleotides within "CpG islands." Among these, DNA methylation is uniquely reversible, and changes in methylation state are known to affect gene expression (*e.g.*, transcriptional initiation of genes where CpG islands located at or near the promoter region) or genomic stability.

25 *Methylation of CpG dinucleotides within CpG islands.* DNA, in higher order eukaryotic organisms, is methylated only at cytosine residues located 5' to guanosine residues in CpG dinucleotides. This covalent modification of the C-5 position of the cytosine base by the enzyme DNA (cytosine-5)-methyltransferase results in the formation of 5-methylcytosine (5-mCyt), and gives this base unique properties (*e.g.*, susceptibility to
30 undergo spontaneous deamination). This enzymatic conversion is the only epigenetic modification of DNA known to exist in vertebrates, and is essential for normal embryonic development (Bird, A.P., *Cell* 70:5-8, 1992; Laird & Jaenisch, *Human Molecular Genetics* 3:1487-1495, 1994; Li et al., *Cell* 69:915-926, 1992).

The presence of 5-mCyt at CpG dinucleotides has resulted in the 5-fold depletion of
35 this sequence in the genome during the course of vertebrate evolution (Schroderet & Gartler, *Proc. Nat. Acad. Sci. USA* 89:957-961, 1992), presumably due to spontaneous deamination of 5-mCyt to Thymidine. Certain areas of the genome, however, do not show such depletion,

and are referred to as "CpG islands" (Bird, A.P., *Nature* 321:209-213, 1986; Gardiner-Garden & Frommer, *J. Mol. Biol.* 196:261-282, 1987). These CpG islands comprise only approximately 1% of the vertebrate genome, yet account for about 15% of the total number of genomic CpG dinucleotides (Antequera & Bird, *Proc. Nat. Acad. Sci. USA* 90:11995-11999, 1993). CpG islands contain the expected (*i.e.*, the non-evolutionarily depleted) frequency of CpGs (with an Observed/Expected Ratio¹ >0.6), are GC-rich (with a GC Content² >0.5) and are typically between about 0.2 to about 1 kb in length.

Methylation within CpG islands affects gene expression. CpG islands are located upstream of many housekeeping and tissue-specific genes, but may also extend into gene coding regions (Cross & Bird, *Current Opinions in Genetics and Development* 5:309-314, 1995; Larsen et al., *Genomics* 13:1095-1107, 1992). The methylation of cytosines within CpG islands in somatic tissues is believed to affect gene expression. Methylation has been inversely correlated with gene activity and may lead to decreased gene expression by a variety of mechanisms including inhibition of transcription initiation (Bird, A.P., *Nature* 321:209-213, 1986; Delgado et al., *EMBO Journal* 17:2426-2435, 1998), disruption of local chromatin structure (Counts & Goodman, *Molecular Carcinogenesis* 11:185-188, 1994; Antequera et al., *Cell* 62:503-514, 1990), and recruitment of proteins that interact specifically with methylated sequences and thereby directly or indirectly prevent transcription factor binding (Bird, A.P., *Cell* 70:5-8, 1992; Counts & Goodman, *Molecular Carcinogenesis* 11:185-188, 1994; Cedar, H., *Cell* 53:3-4, 1988). Many studies have demonstrated the effect of methylation of CpG islands on gene expression (*e.g.*, the *CDKN2A/p16* gene; Gonzalez-Zulueta et al., *Cancer Research* 55:4531-4535, 1995), but most CpG islands on autosomal genes remain unmethylated in the germline, and methylation of these islands is usually independent of gene expression. Tissue-specific genes are typically unmethylated in the respective target organs but are methylated in the germline and in non-expressing adult tissues, while CpG islands of constitutively expressed housekeeping genes are normally unmethylated in the germline and in somatic tissues.

Methylation within CpG islands affects the expression of genes involved in cancer. Data from a group of studies show the presence of altered methylation in cancer cells relative to non-cancerous cells. These studies show not only alteration of the overall genomic levels of DNA methylation, but also changes in the distribution of methyl groups. For example, abnormal methylation of CpG islands that are associated with tumor suppressor genes or oncogenes within a cell may cause altered gene expression. Such altered gene expression may provide a population of cells with a selective growth advantage and thereby result in selection of these cells to the detriment of the organism (*i.e.*, cancer).

¹ Calculated as: [number of CpG sites / (number of C bases X number of G bases)] X band length for each fragment.

² Calculated as: (number of C bases + number of G bases) / band length for each fragment.

Insufficient correlative data. Unfortunately, the mere knowledge of the basic existence of altered methylation of CpG dinucleotides within CpG islands of cancer cells relative to normal cells, or of the fact that in particular instances such methylation changes result in altered gene expression (or chromatin structure or stability), is inadequate to allow for effective diagnostic, prognostic and therapeutic application of this knowledge. This is because only a limited number of CpG islands have been characterized, and thus there is insufficient knowledge, as to which particular CpG islands, among many, are actually involved in, or show significant correlation with cancer or the etiology thereof. Moreover, complex methylation patterns, involving a plurality of methylation-altered DNA sequences, including those that may have the sequence composition to qualify as CpG islands, may exist in particular cancers.

Therefore there is a need in the art to identify and characterize specific methylation altered DNA sequences, and to correlate them with cancer to allow for their diagnostic, prognostic and therapeutic application.

Summary of the Invention

The present invention provides for a diagnostic or prognostic assay for cancer, comprising: obtaining a tissue sample from a test tissue; performing a methylation assay on DNA derived from the tissue sample, wherein the methylation assay determines the methylation state of a CpG dinucleotide within a DNA sequence of the DNA, and wherein the DNA sequence is a sequence selected from the group consisting of sequences of SEQ ID NOS:1-103, sequences having a nucleotide sequence at least 90% identical to sequences of SEQ ID NOS:1-103, CpG island sequences associated with sequences of SEQ ID NOS:1-103, CpG island sequences associated with sequences having a nucleotide sequence at least 90% identical to sequences of SEQ ID NOS:1-103, and combinations thereof, wherein the CpG island sequence associated with the sequence of the particular SEQ ID NO is that contiguous sequence of genomic DNA that encompasses at least one nucleotide of the particular SEQ ID NO sequence, and satisfies the criteria of having both a frequency of CpG dinucleotides corresponding to an Observed/Expected Ratio >0.6 , and a GC Content >0.5 ; and determining a diagnosis or prognosis based, at least in part, upon the methylation state of the CpG dinucleotide within the DNA sequence. Preferably, the DNA sequence is a sequence selected from the group consisting of CpG island sequences associated with sequences of SEQ ID NOS:1-103, CpG island sequences associated with sequences having a nucleotide sequence at least 90% identical to sequences of SEQ ID NOS:1-103, and combinations thereof. Preferably, the DNA sequence is a sequence selected from the group consisting of CpG island sequences associated with sequences of SEQ ID NOS: 2, 4, 6, 7, 9-16, 19, 20, 22-33, 35-43, 48, 51-55, 59, 60, 64, 71, 76, 78-81, 84 and 87-90, and combinations thereof. Preferably, the methylation assay procedure is selected from the group

consisting of MethyLight, MS-SnuPE (methylation-sensitive single nucleotide primer extension), MSP (methylation-specific PCR), MCA (methylated CpG island amplification), COBRA (combined bisulfite restriction analysis), and combinations thereof. Preferably, the methylation state of the CpG dinucleotide within the DNA sequence is that of

- 5 hypermethylation, hypomethylation or normal methylation. Preferably, the cancer is selected from the group consisting of bladder cancer, prostate cancer, colon cancer, lung cancer, renal cancer, leukemia, breast cancer, uterine cancer, astrocytoma, glioblastoma, and neuroblastoma. Preferably, the cancer is bladder cancer, or prostate cancer.

- 10 The present invention further provides a kit useful for the detection of a methylated CpG-containing nucleic acid comprising a carrier means containing one or more containers comprising: a container containing a probe or primer which hybridizes to any region of a sequence selected from the group consisting of SEQ ID NOS:1-103, and sequences having a nucleotide sequence at least 90% identical to sequences of SEQ ID NOS:1-103; and additional standard methylation assay reagents required to affect detection of methylated
- 15 CpG-containing nucleic acid based on the probe or primer. Preferably, the additional standard methylation assay reagents are standard reagents for performing a methylation assay from the group consisting of MethyLight, MS-SNuPE, MSP, MCA, COBRA, and combinations thereof. Preferably, the probe or primer comprises at least about 12 to 15 nucleotides of a sequence selected from the group consisting of SEQ ID NOS:1-103, and
- 20 sequences having a nucleotide sequence at least 90% identical to sequences of SEQ ID NOS:1-103.

- The present invention further provides an isolated nucleic acid molecule comprising a methylated or unmethylated polynucleotide sequence selected from the group consisting of
- 25 SEQ ID NO:1, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:18, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:32, SEQ ID NO:34, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:42, SEQ ID NO:44, SEQ ID NO:51, SEQ ID NO:52, SEQ ID NO:62, SEQ ID NO:64, SEQ ID NO:65, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:74, SEQ ID NO:76, SEQ ID NO:82, SEQ ID NO:83, SEQ ID NO:84, SEQ ID NO:86, SEQ ID NO:90, SEQ ID NO:92, SEQ ID NO:97, and SEQ ID NO:100. Preferably the nucleic acid is methylated. Preferably, the nucleic acid
- 30 is unmethylated.

Detailed Description of the Invention

35 Definitions:

"GC Content" refers, within a particular DNA sequence, to the [(number of C bases + number of G bases) / band length for each fragment].

"Observed/Expected Ratio" ("O/E Ratio") refers to the frequency of CpG

dinucleotides within a particular DNA sequence, and corresponds to the [number of CpG sites / (number of C bases X number of G bases)] X band length for each fragment.

“CpG Island” refers to a contiguous region of genomic DNA that satisfies the criteria of (1) having a frequency of CpG dinucleotides corresponding to an

5 **“Observed/Expected Ratio”** >0.6), and (2) having a **“GC Content”** >0.5. CpG islands are typically, but not always, between about 0.2 to about 1 kb in length. A CpG island sequence associated with a particular SEQ ID NO sequence of the present invention is that contiguous sequence of genomic DNA that encompasses at least one nucleotide of the particular SEQ ID NO sequence, and satisfies the criteria of having both a frequency of CpG dinucleotides
10 corresponding to an Observed/Expected Ratio >0.6), and a GC Content >0.5.

“Methylation state” refers to the presence or absence of 5-methylcytosine (“5-mCyt”) at one or a plurality of CpG dinucleotides within a DNA sequence.

“Hypermethylation” refers to the methylation state corresponding to an *increased* presence of 5-mCyt at one or a plurality of CpG dinucleotides within a DNA sequence of a
15 test DNA sample, relative to the amount of 5-mCyt found at corresponding CpG dinucleotides within a normal control DNA sample.

“Hypomethylation” refers to the methylation state corresponding to a *decreased* presence of 5-mCyt at one or a plurality of CpG dinucleotides within a DNA sequence of a test DNA sample, relative to the amount of 5-mCyt found at corresponding CpG
20 dinucleotides within a normal control DNA sample.

“Methylation assay” refers to any assay for determining the methylation state of a CpG dinucleotide within a sequence of DNA.

“MS.AP-PCR” (Methylation-Sensitive Arbitrarily-Primed Polymerase Chain Reaction) refers to the art-recognized technology that allows for a global scan of the genome
25 using CG-rich primers to focus on the regions most likely to contain CpG dinucleotides, and described by Gonzalzo et al., *Cancer Research* 57:594-599, 1997.

“MethyLight” refers to the art-recognized fluorescence-based real-time PCR technique described by Eads et al., *Cancer Res.* 59:2302-2306, 1999.

“Ms-SNuPE” (Methylation-sensitive Single Nucleotide Primer Extension) refers to
30 the art-recognized assay described by Gonzalzo & Jones, *Nucleic Acids Res.* 25:2529-2531, 1997.

“MSP” (Methylation-specific PCR) refers to the art-recognized methylation assay described by Herman et al. *Proc. Natl. Acad. Sci. USA* 93:9821-9826, 1996, and by US Patent No. 5,786,146.

35 **“COBRA”** (Combined Bisulfite Restriction Analysis) refers to the art-recognized methylation assay described by Xiong & Laird, *Nucleic Acids Res.* 25:2532-2534, 1997.

“MCA” (Methylated CpG Island Amplification) refers to the methylation assay described by Toyota et al., *Cancer Res.* 59:2307-12, 1999, and in WO 00/26401A1.

Overview

The present invention provides for 103 DNA sequences (*i.e.*, “marker sequences”) having distinct methylation patterns in cancer, as compared to normal tissue. These methylation-altered DNA sequence embodiments correspond to 103 DNA fragments isolated from bladder and prostate cancer patients, and in many instances, represent novel sequences not found in the GenBank database. *None* of the instant sequence embodiments have previously been characterized with respect to their methylation pattern in human cancers including, but not limited to, those of the bladder and prostate. The significance of such methylation patterns lies in the value of altered fragments as potential prognostic, diagnostic and therapeutic markers in the treatment of human cancers.

Identification of Methylation-altered Marker Sequences in Genomic DNA

The MS.AP-PCR technique was used to scan the genomes of bladder or prostate cancer patients for DNA methylation changes relative to normal individuals, because the pattern is known to be highly conserved. A total of 103 DNA sequence embodiments (methylation-altered DNA sequences; “marker sequences”) were isolated and characterized as having distinct methylation patterns in cancer, as compared to normal tissue.

Methods for the Identification of Marker Sequences in Genomic DNA. There are a variety of art-recognized genome scanning methods that have been used to identify altered methylation sites in cancer cells. For example, one method involves restriction landmark genomic scanning (Kawai et al., *Mol. Cell. Biol.* 14:7421-7427, 1994), another involves MCA (methylated CpG island amplification; Toyota et al., *Cancer Res.* 59:2307-12, 1999), and yet another involves MS.AP-PCR (Methylation-Sensitive Arbitrarily-Primed Polymerase Chain Reaction; Gonzalgo et al., *Cancer Res.* 57:594-599, 1997), which allows for a global scan of the genome using CG-rich primers to focus on the regions most likely to contain CpG dinucleotides. The MS.AP-PCR technique used in the present invention is a rapid and efficient method to screen (“scan”) for altered methylation patterns in genomic DNA and to isolate specific sequences associated with these changes.

Briefly, genomic DNA from the tissue of bladder or prostate cancer patients was prepared using standard, art-recognized methods. Restriction enzymes (*e.g.*, HpaII) with different sensitivities to cytosine methylation in their recognition sites were used to digest these genomic DNAs prior to arbitrarily primed PCR amplification with GC-rich primers. Fragments that showed differential methylation (*e.g.*, *hypermethylation* or *hypomethylation*, based on the methylation sensitivity of the restriction enzyme, or upon DNA sequence analysis or Ms-SNuPE analysis; Gonzalgo & Jones, *Nucleic Acids Res* 25:2529-2531, 1997) were cloned and sequenced after resolving the PCR products on high-resolution polyacrylamide gels. The cloned fragments were used as probes for Southern blot analysis to

confirm differential methylation of these regions in the tissue. Methods for DNA cloning, sequencing, PCR, high-resolution polyacrylamide gel resolution and Southern blot analysis are well known by those of ordinary skill in the relevant art.

Results. A total of 500 DNA fragments that underwent either hypermethylation (an increase in the level of methylation relative to normal) or hypomethylation (a decrease in the level of methylation relative to normal) were isolated from the scanned patients genomic DNA. A total of 178 of these fragments were sequenced, of which 103 were *novel* in that they corresponded to DNA loci whose methylation pattern had not previously been characterized. The corresponding sequences are disclosed as [SEQ ID NOS:1-103], wherein for certain sequences, the letter "n" refers to an undetermined nucleotide base.

Novel marker sequences identified by MS.AP-PCR. Table I shows an overall summary of methylation patterns and sequence data corresponding to the 103 DNA fragments identified by MS.AP-PCR. A total of 103 fragments were sequenced following identification as becoming either hypermethylated (gain of methylation; noted as having a hypermethylation pattern) or hypomethylated (loss of methylation; noted as having a hypomethylation pattern) relative to normal tissue. For the fragments of each category, the "Average GC Content" is shown, calculated as (number of C bases + number of G bases)/band length for each fragment, as well as the average Observed/Expected Ratio ("O/E Ratio"), calculated as [number of CpG sites/(number of C bases X number of G bases)] X band length for each fragment. Additionally, the percent of fragments that qualify as CpG islands is listed, and corresponds to the percentage of all fragments within each category that have sequence compositions that satisfy the criteria of having a "GC Content" >0.5 and an "O/E Ratio" >0.6.

Thus, of these 103 fragments identified by MS.AP-PCR, 60 showed hypermethylation (Table I, upper row; Table II, [SEQ ID NOS:1-60]) while 43 showed hypomethylation (Table I, lower row; Table II, [SEQ ID NOS:61-103]). Moreover, 55 (43 hypermethylated, and 12 hypomethylated) of the 103 fragments correspond to CpG islands (*i.e.*, fulfill the criteria of a GC content >0.5 and an Observed/Expected Ratio >0.6;), whereas the other 48 (17 hypermethylated and 31 hypomethylated) fragments do not meet the criteria for CpG islands (*see* Table II).

TABLE I. Summary of 103 DNA Fragments Identified by MS.AP-PCR

DNA Fragment Type	Methylation Pattern (relative to normal)	Number of Fragments (103 total)	Average GC Content	Average O/E Ratio	Percent that correspond to CpG Islands
Hypermethylated Fragments	Hyper-methylation	60	0.54	0.72	72%
Hypomethylated Fragments	Hypo-methylation	43	0.52	0.48	28%

Table II shows a summary of methylation pattern and sequence data for each individual sequence embodiment ([SEQ ID NOS:1-103]), corresponding to the 103 DNA fragments identified by MS.AP-PCR. Data for the 103 fragments was divided into either hypermethylated ([SEQ ID NOS:1-60]) or hypomethylated ([SEQ ID NOS:61-103]) categories. Table II also lists, for each sequence embodiment, the corresponding "Fragment Name," fragment "Size" (in base pairs; "bp"), "GC Content," Observed/Expected Ratio ("O/E Ratio"), "Description" (*i.e.*, as a CpG island if criteria are met), "Inventor Initials" (IDCM = Isabel D.C. Markl, JC = Jonathan Cheng, GL = Gangning Liang, HF = Hualin Fu, YT = Yoshitaka Tomigahara), "Cancer Source," and "Chromosome Match" to the GenBank database. A dash ("-") indicates that no GenBank chromosome match existed, or that only a low-scoring partial match was found. Averages of the "GC Content" and "O/E Ratio," along with the percent of fragments that are CpG islands, are listed after the last member of both the hypermethylated and hypomethylated categories.

Therefore, the present invention provides for 103 DNA fragments and corresponding marker sequence embodiments (*i.e.*, methylation-altered DNA sequences) that are useful in cancer prognostic, diagnostic and therapeutic applications.

Additionally, at least 55 of these 103 sequences correspond to CpG islands (based on GC Content and O/E ration); namely [SEQ ID NOS:2, 4, 6, 7, 9-16, 19, 20, 22-33, 35-43, 48, 51-55, 59, 60, 64, 71, 76, 78-81, 84 and 87-90]. Thus, based on the fact that the methylation state of a portion of a given CpG island is generally representative of the island as a whole, the present invention further encompassed the novel use of the 55 CpG islands associated with [SEQ ID NOS:2, 4, 6, 7, 9-16, 19, 20, 22-33, 35-43, 48, 51-55, 59, 60, 64, 71, 76, 78-81, 84 and 87-90] in cancer prognostic, diagnostic and therapeutic applications, where a CpG island sequence associated with the sequence of a particular SEQ ID NO is that contiguous sequence of genomic DNA that encompasses at least one nucleotide of the particular SEQ ID NO sequence, and satisfies the criteria of having both a frequency of CpG dinucleotides corresponding to an Observed/Expected Ratio >0.6, and a GC Content >0.5.

TABLE II. Summary of MS.AP-PCR Fragments Sequenced

Methylation Pattern	Fragment Name	Size (bp)	GC Content	O/E Ratio	Description	Inventor Initials	Cancer Source	Chromosome Matches	[SEQ ID NO]
Hyper-methylation Category	11-1A	510	0.44	0.74		IDCM	Bladder	-	1
	14-3B	313	0.58	0.74	CpG Island	IDCM	Bladder	2	2
	18-2B	165	0.57	0.45		IDCM	Bladder	7	3
	24-1B	601	0.51	0.72	CpG Island	IDCM	Bladder	Xp11	4
	26-1B	801	0.48	0.56		IDCM	Bladder	-	5
	26-2C	204	0.50	0.63	CpG Island	IDCM	Bladder	-	6
	30-3D	205	0.55	1.25	CpG Island	IDCM	Bladder	14	7
	32-3E	597	0.57	0.10		IDCM	Bladder	20q12-13.1	8
	34-2B	500	0.62	0.66	CpG Island	IDCM	Bladder	20	9
	34-4B	343	0.70	0.81	CpG Island	IDCM	Bladder	-	10

Methylation Pattern	Fragment Name	Size (bp)	GC Content	O/E Ratio	Description	Inventor Initials	Cancer Source	Chromosome Matches	[SEQ ID NO]
	34-5D	291	0.62	0.96	CpG Island	IDCM	Bladder	9	11
	34-6A	266	0.64	0.93	CpG Island	IDCM	Bladder	-	12
	35-1C	553	0.64	0.63	CpG Island	IDCM	Bladder	-	13
	36-2D	156	0.60	0.58	CpG Island	IDCM	Bladder	10	14
	38-1A	300	0.70	0.80	CpG Island	IDCM	Bladder	10	15
	38-2B	196	0.56	0.89	CpG Island	IDCM	Bladder	15	16
	7-8E	299	0.59	0.39		IDCM	Bladder	17q21-22	17
	83-4B	363	0.54	0.49		IDCM	Bladder	-	18
	84-1D	322	0.55	0.90	CpG Island	IDCM	Bladder	7	19
	101-3E	255	0.57	0.83	CpG Island	IDCM	Bladder	17	20
	M1-5A	406	0.45	0.96		IDCM	Bladder	1	21
	U2-8E	210	0.56	0.61	CpG Island	IDCM	Bladder	2	22
	U12-1A	310	0.56	0.81	CpG Island	IDCM	Bladder	2	23
	U7-4A	305	0.59	0.80	CpG Island	IDCM	Bladder	-	24
	NU9-5A	379	0.67	0.83	CpG Island	JC	Bladder	-	25
	3-17-8-B	625	0.48	0.72	CpG Island	GL	Bladder	18	26
	4-10-4-A	499	0.55	0.30	CpG Island	GL	Bladder	7	27
	1-1-1-A	561	0.58	0.98	CpG Island	GL	Bladder	20	28
	3-17-8-A	717	0.50	0.68	CpG Island	GL	Bladder	17	29
	G145-H	280	0.50	1.10	CpG Island	GL	Bladder	11	30
	1-1-1-D	270	0.50	0.60	CpG Island	GL	Bladder	2	31
	1-1-1-C	347	0.65	1.25	CpG Island	GL	Bladder	-	32
	G178-A	342	0.55	0.85	CpG Island	GL	Bladder	2	33
	34-A'	370	0.62	0.44		HF	Prostate	-	34
	34-D	213	0.53	0.74	CpG Island	HF	Prostate	2	35
	35-D	173	0.56	0.66	CpG Island	HF	Prostate	3	36
	36-A	369	0.67	0.70	CpG Island	HF	Prostate	-	37
	40-A	123	0.60	1.16	CpG Island	HF	Prostate	-	38
	91-1	450	0.64	0.86	CpG Island	YT	Bladder	5 or 16q24.3	39
	93-2	593	0.51	0.68	CpG Island	YT	Bladder	Xp11	40
	93-3	457	0.52	0.94	CpG Island	YT	Bladder	Xp22.1-22.3	41
	94-8	211	0.66	0.96	CpG Island	YT	Bladder	-	42
	95-5	141	0.63	0.79	CpG Island	YT	Bladder	14	43
	97-5	559	0.56	0.40		YT	Bladder	-	44
	98-1	433	0.46	0.96		YT	Bladder	1	45
	100-1	487	0.59	0.58		YT	Bladder	14	46
	100-2	403	0.60	0.47		YT	Bladder	3	47
	100-6	155	0.57	0.99	CpG Island	YT	Bladder	20	48
	4-2	256	0.57	0.40		YT	Bladder	7	49
	5-8	224	0.47	0.96		YT	Bladder	5	50
	6-4	313	0.70	0.82	CpG Island	YT	Bladder	-	51
	7-6	385	0.70	0.88	CpG Island	YT	Bladder	-	52
	13-3	307	0.59	0.89	CpG Island	YT	Bladder	10	53
	15-2	182	0.62	0.92	CpG Island	YT	Bladder	13	54
	23-2	523	0.54	0.87	CpG Island	YT	Bladder	Xp22.1-22.3	55
	39-2	795	0.46	0.64		YT	Bladder	13	56
	40-2	438	0.62	0.51		YT	Bladder	10	57
	41-3	611	0.47	0.70		YT	Bladder	18	58
	105-4	291	0.58	0.71	CpG Island	YT	Bladder	5	59
	107-8	226	0.53	0.96	CpG Island	YT	Bladder	11	60
AVERAGE			0.54	0.72	72% islands				
Hypo-methylation Category	14-2B	580	0.55	0.51		IDCM	Bladder	2	61
	16-1B	633	0.56	0.39		IDCM	Bladder	-	62
	18-1B	703	0.45	0.35		IDCM	Bladder	17	63

Methylation Pattern	Fragment Name	Size (bp)	GC Content	O/E Ratio	Description	Inventor Initials	Cancer Source	Chromosome Matches	[SEQ ID NO]
	19-1B	420	0.66	0.87	CpG Island	IDCM	Bladder	-	64
	20-1B	496	0.61	0.59		IDCM	Bladder	-	65
	21-2C	637	0.60	0.33		IDCM	Bladder	9q34	66
	29-1A	595	0.55	0.27		IDCM	Bladder	Xp11.23	67
	29-2B	580	0.47	0.77		IDCM	Bladder	-	68
	32-1A	589	0.59	0.48		IDCM	Bladder	-	69
	34-1B	450	0.42	0.46		IDCM	Bladder	-	70
	34-3B	432	0.70	0.61	CpG Island	IDCM	Bladder	-	71
	32-2B	748	0.47	0.24		IDCM	Bladder	2	72
	32-4B	599	0.57	0.15		IDCM	Bladder	20q12-13.1	73
	32-5B	614	0.58	0.20		IDCM	Bladder	-	74
	33-1A	552	0.54	0.32		IDCM	Bladder	10	75
	5-1E	501	0.61	1.04	CpG Island	IDCM	Bladder	-	76
	6-1A	826	0.55	0.36		IDCM	Bladder	22q13.32-13.33	77
	7-5D	433	0.59	0.85	CpG Island	IDCM	Bladder	5	78
	8-7C	424	0.58	0.83	CpG Island	IDCM	Bladder	5	79
	30-6D	285	0.63	0.72	CpG Island	IDCM	Bladder	1	80
	66-2E	401	0.54	0.82	CpG Island	IDCM	Bladder	16	81
	78-1C	268	0.54	0.41		IDCM	Bladder	-	82
	97-2E	989	0.53	0.16		IDCM	Bladder	-	83
	M1-8C	250	0.64	0.99	CpG Island	IDCM	Bladder	-	84
	M2-5A	402	0.50	0.45		IDCM	Bladder	5	85
	M1-4P	595	0.43	0.41		IDCM	Bladder	-	86
	M12-10A	304	0.53	0.76	CpG Island	IDCM	Bladder	7	87
	M12-12C	296	0.51	0.64	CpG Island	IDCM	Bladder	17	88
	M2-8M	220	0.67	0.62	CpG Island	IDCM	Bladder	6q27	89
	NU4-3A	273	0.63	1.02	CpG Island	JC	Bladder	-	90
	NU5-2A	361	0.44	0.73		JC	Bladder	6q14.3-15	91
	88-5	462	0.62	0.39		YT	Bladder	-	92
	90-1	591	0.66	0.45		YT	Bladder	19	93
	91-3	279	0.58	0.45		YT	Bladder	5 or 16q24.3	94
	91-4	351	0.55	0.30		YT	Bladder	18q23	95
	91-7	171	0.61	0.59		YT	Bladder	11	96
	89-3	743	0.55	0.43		YT	Bladder	-	97
	94-2	589	0.53	0.41		YT	Bladder	22q13.31-13.32	98
	94-3	538	0.53	0.49		YT	Bladder	5 or 18	99
	94-4	486	0.61	0.57		YT	Bladder	-	100
	94-5	450	0.60	0.45		YT	Bladder	1p36.2-36.3	101
	94-6	292	0.58	0.32		YT	Bladder	8 or 9	102
	96-4	395	0.63	0.54		YT	Bladder	9	103
AVERAGE			0.52	0.48	28% islands				

Diagnostic and Prognostic Assays for Cancer. The present invention provides for diagnostic and prognostic cancer assays based on determination of the methylation state of one or more of the disclosed 103 methylation-altered DNA sequence embodiments. Typically, such assays involve obtaining a tissue sample from a test tissue, performing a methylation assay on DNA derived from the tissue sample, and making a diagnosis or prognosis based thereon.

The methylation assay is used to determine the methylation state of one or a plurality of CpG dinucleotide within a DNA sequence of the DNA sample. According to the present invention, possible methylation states include *hypermethylation* and *hypomethylation*, relative to a normal state (*i.e.*, non-cancerous control state). Hypermethylation and hypomethylation refer to the methylation states corresponding to an *increased* or *decreased*, respectively, presence 5-methylcytosine ("5-mCyt") at one or a plurality of CpG dinucleotides within a DNA sequence of the test sample, relative to the amount of 5-mCyt found at corresponding CpG dinucleotides within a normal control DNA sample.

A diagnosis or prognosis is based, at least in part, upon the determined methylation state of the sample DNA sequence compared to control data obtained from normal, non-cancerous tissue.

Methylation Assay Procedures. Various methylation assay procedures are known in the art, and can be used in conjunction with the present invention. These assays allow for determination of the methylation state of one or a plurality of CpG dinucleotides (*e.g.*, CpG islands) within a DNA sequence. Such assays involve, among other techniques, DNA sequencing of bisulfite-treated DNA, PCR (for sequence-specific amplification), Southern blot analysis, use of methylation-sensitive restriction enzymes, etc.

For example, genomic sequencing has been simplified for analysis of DNA methylation patterns and 5-methylcytosine distribution by using bisulfite treatment (Frommer et al., *Proc. Natl. Acad. Sci. USA* 89:1827-1831, 1992). Additionally, restriction enzyme digestion of PCR products amplified from bisulfite-converted DNA is used, *e.g.*, the method described by Sadri & Hornsby (*Nucl. Acids Res.* 24:5058-5059, 1996), or COBRA (Combined Bisulfite Restriction Analysis) (Xiong & Laird, *Nucleic Acids Res.* 25:2532-2534, 1997).

COBRA. COBRA analysis is a quantitative methylation assay useful for determining DNA methylation levels at specific gene loci in small amounts of genomic DNA (Xiong & Laird, *Nucleic Acids Res.* 25:2532-2534, 1997). Briefly, restriction enzyme digestion is used to reveal methylation-dependent sequence differences in PCR products of sodium bisulfite-treated DNA. Methylation-dependent sequence differences are first introduced into the genomic DNA by standard bisulfite treatment according to the procedure described by Frommer et al. (*Proc. Natl. Acad. Sci. USA* 89:1827-1831, 1992). PCR amplification of the bisulfite converted DNA is then performed using primers specific for the interested CpG islands, followed by restriction endonuclease digestion, gel electrophoresis, and detection using specific, labeled hybridization probes. Methylation levels in the original DNA sample are represented by the relative amounts of digested and undigested PCR product in a linearly quantitative fashion across a wide spectrum of DNA methylation levels. In addition, this technique can be reliably applied to DNA obtained from microdissected paraffin-embedded tissue samples. Typical reagents (*e.g.*, as might be found in a typical COBRA-based kit) for

COBRA analysis may include, but are not limited to: PCR primers for specific gene (or methylation-altered DNA sequence or CpG island); restriction enzyme and appropriate buffer; gene-hybridization oligo; control hybridization oligo; kinase labeling kit for oligo probe; and radioactive nucleotides. Additionally, bisulfite conversion reagents may include:
5 DNA denaturation buffer; sulfonation buffer; DNA recovery reagents or kit (e.g., precipitation, ultrafiltration, affinity column); desulfonation buffer; and DNA recovery components.

Preferably, assays such as "MethyLight" (a fluorescence-based real-time PCR technique) (Eads et al., *Cancer Res.* 59:2302-2306, 1999), Ms-SNuPE (Methylation-sensitive
10 Single Nucleotide Primer Extension) reactions (Gonzalzo & Jones, *Nucleic Acids Res.* 25:2529-2531, 1997), methylation-specific PCR ("MSP"; Herman et al., *Proc. Natl. Acad. Sci. USA* 93:9821-9826, 1996; US Patent No. 5,786,146), and methylated CpG island amplification ("MCA"; Toyota et al., *Cancer Res.* 59:2307-12, 1999) are used alone or in combination with other of these methods.

15 **MethyLight.** The MethyLight assay is a high-throughput quantitative methylation assay that utilizes fluorescence-based real-time PCR (TaqMan®) technology that requires no further manipulations after the PCR step (Eads et al., *Cancer Res.* 59:2302-2306, 1999). Briefly, the MethyLight process begins with a mixed sample of genomic DNA that is converted, in a sodium bisulfite reaction, to a mixed pool of methylation-dependent sequence
20 differences according to standard procedures (the bisulfite process converts unmethylated cytosine residues to uracil). Fluorescence-based PCR is then performed either in an "unbiased" (with primers that do not overlap known CpG methylation sites) PCR reaction, or in a "biased" (with PCR primers that overlap known CpG dinucleotides) reaction. Sequence discrimination can occur either at the level of the amplification process or at the level of the
25 fluorescence detection process, or both.

The MethyLight may assay be used as a quantitative test for methylation patterns in the genomic DNA sample, wherein sequence discrimination occurs at the level of probe hybridization. In this quantitative version, the PCR reaction provides for unbiased amplification in the presence of a fluorescent probe that overlaps a particular putative
30 methylation site. An unbiased control for the amount of input DNA is provided by a reaction in which neither the primers, nor the probe overlap any CpG dinucleotides. Alternatively, a qualitative test for genomic methylation is achieved by probing of the biased PCR pool with either control oligonucleotides that do not "cover" known methylation sites (a fluorescence-based version of the "MSP" technique), or with oligonucleotides covering potential
35 methylation sites.

The MethyLight process can be used with a "TaqMan®" probe in the amplification process. For example, double-stranded genomic DNA is treated with sodium bisulfite and subjected to one of two sets of PCR reactions using TaqMan® probes; e.g., with either

biased primers and TaqMan® probe, or unbiased primers and TaqMan® probe. The TaqMan® probe is dual-labeled with fluorescent “reporter” and “quencher” molecules, and is designed to be specific for a relatively high GC content region so that it melts out at about 10 °C higher temperature in the PCR cycle than the forward or reverse primers. This allows the TaqMan® probe to remain fully hybridized during the PCR annealing/extension step. As the Taq polymerase enzymatically synthesizes a new strand during PCR, it will eventually reach the annealed TaqMan® probe. The Taq polymerase 5’ to 3’ endonuclease activity will then displace the TaqMan® probe by digesting it to release the fluorescent reporter molecule for quantitative detection of its now unquenched signal using a real-time fluorescent detection system.

Typical reagents (*e.g.*, as might be found in a typical MethyLight-based kit) for MethyLight analysis may include, but are not limited to: PCR primers for specific gene (or methylation-altered DNA sequence or CpG island); TaqMan® probes; optimized PCR buffers and deoxynucleotides; and Taq polymerase.

Ms-SNuPE. The Ms-SNuPE technique is a quantitative method for assessing methylation differences at specific CpG sites based on bisulfite treatment of DNA, followed by single-nucleotide primer extension (Gonzalzo & Jones, *Nucleic Acids Res.* 25:2529-2531, 1997). Briefly, genomic DNA is reacted with sodium bisulfite to convert unmethylated cytosine to uracil while leaving 5-methylcytosine unchanged. Amplification of the desired target sequence is then performed using PCR primers specific for bisulfite-converted DNA, and the resulting product is isolated and used as a template for methylation analysis at the CpG site(s) of interest. Small amounts of DNA can be analyzed (*e.g.*, microdissected pathology sections), and it avoids utilization of restriction enzymes for determining the methylation status at CpG sites. Typical reagents (*e.g.*, as might be found in a typical Ms-SNuPE-based kit) for Ms-SNuPE analysis may include, but are not limited to: PCR primers for specific gene (or methylation-altered DNA sequence or CpG island); optimized PCR buffers and deoxynucleotides; gel extraction kit; positive control primers; Ms-SNuPE primers for specific gene; reaction buffer (for the Ms-SNuPE reaction); and radioactive nucleotides. Additionally, bisulfite conversion reagents may include: DNA denaturation buffer; sulfonation buffer; DNA recovery reagents or kit (*e.g.*, precipitation, ultrafiltration, affinity column); desulfonation buffer; and DNA recovery components.

MSP. MSP (methylation-specific PCR) allows for assessing the methylation status of virtually any group of CpG sites within a CpG island, independent of the use of methylation-sensitive restriction enzymes (Herman et al. *Proc. Natl. Acad. Sci. USA* 93:9821-9826, 1996; US Patent No. 5,786,146). Briefly, DNA is modified by sodium bisulfite converting all unmethylated, but not methylated cytosines to uracil, and subsequently amplified with primers specific for methylated versus unmethylated DNA. MSP requires only small quantities of DNA, is sensitive to 0.1% methylated alleles of a given CpG island locus, and

can be performed on DNA extracted from paraffin-embedded samples. Typical reagents (e.g., as might be found in a typical MSP-based kit) for MSP analysis may include, but are not limited to: methylated and unmethylated PCR primers for specific gene (or methylation-altered DNA sequence or CpG island), optimized PCR buffers and deoxynucleotides, and specific probes.

MCA. The MCA technique is a method that can be used to screen for altered methylation patterns in genomic DNA, and to isolate specific sequences associated with these changes (Toyota et al., *Cancer Res.* 59:2307-12, 1999). Briefly, restriction enzymes with different sensitivities to cytosine methylation in their recognition sites are used to digest genomic DNAs from primary tumors, cell lines, and normal tissues prior to arbitrarily primed PCR amplification. Fragments that show differential methylation are cloned and sequenced after resolving the PCR products on high-resolution polyacrylamide gels. The cloned fragments are then used as probes for Southern analysis to confirm differential methylation of these regions. Typical reagents (e.g., as might be found in a typical MCA -based kit) for MCA analysis may include, but are not limited to: PCR primers for arbitrary priming Genomic DNA; PCR buffers and nucleotides, restriction enzymes and appropriate buffers; gene-hybridization oligos or probes; control hybridization oligos or probes.

Kits for Detection of Methylated CpG-containing Nucleic Acid. The reagents required to perform one or more art-recognized methylation assays (including those identified above) are combined with primers or probes comprising the sequences of SEQ ID NOS:1-103, or portions thereof, to determine the methylation state of CpG-containing nucleic acids. For example, the MethyLight, Ms-SNuPE, MCA, COBRA, and MSP methylation assays could be used alone or in combination, along with primers or probes comprising the sequences of SEQ ID NOS:1-103, or portions thereof, to determine the methylation state of a CpG dinucleotide within a genomic sequence corresponding to SEQ ID NOS:1-103, or to CpG island sequences associated with sequences of SEQ ID NOS:1-103, where the CpG island sequence associated with the sequence of the particular SEQ ID NO is that contiguous sequence of genomic DNA that encompasses at least one nucleotide of the particular SEQ ID NO sequence, and satisfies the criteria of having both a frequency of CpG dinucleotides corresponding to an Observed/Expected Ratio >0.6, and a GC Content >0.5.

We claim:

1. A diagnostic or prognostic assay for cancer, comprising:

(a) obtaining a tissue sample from a test tissue;

(b) performing a methylation assay on DNA derived from the tissue sample,

5 wherein the methylation assay determines the methylation state of a CpG dinucleotide within a DNA sequence of the DNA, and wherein the DNA sequence is a sequence selected from the group consisting of sequences of SEQ ID NOS:1-103, sequences having a nucleotide sequence at least 90% identical to sequences of SEQ ID NOS:1-103, CpG island sequences associated with sequences of SEQ ID NOS:1-103, CpG island sequences associated with
10 sequences having a nucleotide sequence at least 90% identical to sequences of SEQ ID NOS:1-103, and combinations thereof, wherein the CpG island sequence associated with the sequence of the particular SEQ ID NO is that contiguous sequence of genomic DNA that encompasses at least one nucleotide of the particular SEQ ID NO sequence, and satisfies the criteria of having both a frequency of CpG dinucleotides corresponding to an
15 Observed/Expected Ratio >0.6 , and a GC Content >0.5 ; and

(c) determining a diagnosis or prognosis based, at least in part, upon the methylation state of the CpG dinucleotide within the DNA sequence.

2. The diagnostic or prognostic assay of claim 1 wherein the DNA sequence is a sequence selected from the group consisting of CpG island sequences associated with
20 sequences of SEQ ID NOS:1-103, CpG island sequences associated with sequences having a nucleotide sequence at least 90% identical to sequences of SEQ ID NOS:1-103, and combinations thereof.

3. The diagnostic or prognostic assay of claim 2 wherein the DNA sequence is a sequence selected from the group consisting of CpG island sequences associated with
25 sequences of SEQ ID NOS: 2, 4, 6, 7, 9-16, 19, 20, 22-33, 35-43, 48, 51-55, 59, 60, 64, 71, 76, 78-81, 84 and 87-90, and combinations thereof.

4. The diagnostic or prognostic assay of claim 1 wherein the methylation assay procedure is selected from the group consisting of MethyLight, MS-SNuPE, MSP MCA, COBRA, and combinations thereof.

30 5. The diagnostic or prognostic assay of claim 1 wherein the methylation state of the CpG dinucleotide within the DNA sequence is that of hypermethylation, hypomethylation or normal methylation.

6. The diagnostic or prognostic assay of claim 1 wherein the cancer is selected from the group consisting of bladder cancer, prostate cancer, colon cancer, lung cancer, renal
35 cancer, leukemia, breast cancer, uterine cancer, astrocytoma, glioblastoma, and neuroblastoma.

7. A kit useful for the detection of a methylated CpG-containing nucleic acid comprising a carrier means containing one or more containers comprising:

(a) a container containing a probe or primer which hybridizes to any region of a sequence selected from the group consisting of SEQ ID NOS:1-103, and sequences having a nucleotide sequence at least 90% identical to sequences of SEQ ID NOS:1-103; and

5 (b) additional standard methylation assay reagents required to affect detection of methylated CpG-containing nucleic acid based, at least in part, on the probe or primer.

8. The kit of claim 7, wherein the additional standard methylation assay reagents are standard reagents for performing a methylation assay from the group consisting of MethyLight, MS-SNuPE, MSP MCA, COBRA, and combinations thereof.

10 9. The kit of claim 7, wherein the probe or primer comprises at least about 12 to 15 nucleotides of a sequence selected from the group consisting of SEQ ID NOS:1-103, and sequences having a nucleotide sequence at least 90% identical to sequences of SEQ ID NOS:1-103.

10. An isolated nucleic acid molecule comprising a methylated or unmethylated polynucleotide sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:5, 15 SEQ ID NO:6, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:18, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:32, SEQ ID NO:34, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:42, SEQ ID NO:44, SEQ ID NO:51, SEQ ID NO:52, SEQ ID NO:62, SEQ ID NO:64, SEQ ID NO:65, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:74, SEQ ID NO:76, SEQ ID NO:82, SEQ ID NO:83, SEQ ID NO:84, SEQ ID 20 NO:86, SEQ ID NO:90, SEQ ID NO:92, SEQ ID NO:97, and SEQ ID NO:100.

11. The nucleic acid of claim 10, wherein the nucleic acid is methylated.

12. The nucleic acid of claim 10, wherein the nucleic acid is unmethylated.

SEQUENCE LISTING

<110> University of Southern California
Markl, Isabel
Tomigahara, Yoshitaka
Liang, Gangning
Fu, Hualin
Jones, Peter

<120> Methylation Altered DNA Sequences as Markers Associated with Human Cancer

<130> 47465-14

<160> 103

<170> Word V. 97

<210> 1
<211> 510
<212> DNA
<213> Homo sapiens

<220>
<221> unsure
<222> position is 15 nucleotides
<223> "n" refers to an undetermined base

<220>
<221> unsure
<222> position is 28 nucleotides
<223> "n" refers to an undetermined base

<220>
<221> unsure
<222> position is 77 nucleotides
<223> "n" refers to an undetermined base

<220>
<221> unsure
<222> position is 371 nucleotides
<223> "n" refers to an undetermined base

<220>
<221> unsure
<222> position is 399 nucleotides
<223> "n" refers to an undetermined base

<220>
<221> unsure
<222> position is 410 nucleotides
<223> "n" refers to an undetermined base

<220>
<221> unsure
<222> position is 506 nucleotides
<223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 508 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 509 nucleotides
 <223> "n" refers to an undetermined base

<400> 1
 ttgcaagccc ccttngctct tcctttgncc tgcctacat attcagggga tcgcaatctc 60
 actcgcgaaa taatttnttt ctgtaagagg aagccgcctt tccctctcc caccgccaag 120
 gtaaaggctg cttaaagtagc tcttcttgga aggaaaaata ttttaaaaag cagctggggtt 180
 gctctccaca agaagatggc agttttggga aaaccatta tgtgtccaaa tgccgggttc 240
 cttttcttgt ttaacgcttt ttagagggca aaaatgacgc tcatgtgaag cccacaggct 300
 cgagccaatg tcgctgggct aattatgagt ctgcttatcc cactcccaa tatccgagac 360
 gactcactca naagacattt ttactcttcc aagaattgng aattcagaan cagcttcccc 420
 acattctaag agaaaaaaaa acttgtttaa cgggcacggt tttgattttt ttgccgctgg 480
 cgaccttaat taaaagccgg gagctncnna 510

<210> 2
 <211> 313
 <212> DNA
 <213> Homo sapiens

<400> 2
 gcactcttaa aacgcctctc tgcagtccca ggccgcgct cccaagaac tggccagatc 60
 ggcgcgggct tggccctga caactctgcc tctccacct gttgcgttta ctccgtttag 120
 ttggctgtgc agtctctggc ccagggtgtg cttttaaaac tcgaggaacg cgggtgttgg 180
 actcattcgc agcctcttgc ctctgggttc cgtgatecca cggtggcgag ctccaggct 240
 cagcgaggag atctgggttt gaacattcat ctcccatggt actcttttct tgctcctcgc 300
 gtccccaagc cga 313

<210> 3
 <211> 165
 <212> DNA
 <213> Homo sapiens

<400> 3

gcttagcaaa tttccctttt ttattgttgg ttttgtctgt tggtctttac cccctttcct 60
 tttcctgctt cccctgagtc agcaatgctg agcccagcga agcacagggg gccaaaggga 120
 gagacacacg gagcgccccg ggggtcccca gcctcggcgg ccaaa 165

<210> 4
 <211> 601
 <212> DNA
 <213> Homo sapiens

<400> 4
 ggggggagtc gtgcgtgtca gatttaggcc aggaagcga agtcgccagc agcgagagtt 60
 taacctctgt gggcgagag ggttgccggg attcagcgcc cgggaccgtg gatctgtgca 120
 gggagtcata ggtgtgtgtg acatcagtg tggaacattt tggctcgttt tcacaattca 180
 gtcattatcc tttctgcttt cctcctggaa gcattaaggt tgaagttttc ttctaaagat 240
 caaagttttg atttgttata ttagttcgga tttgtttgat tttgtttgt gttcggtttc 300
 aagtgtgat ttgtaacttt tctcccccc caccacacacg ctttttgacc cctgaattat 360
 ttaaaagtcc attgttgag tggcaaaca cctccgagac tcaaaggga aggccatggg 420
 cgctttattc cggtgtgtgc tccaggaacg tgggaaagca gcggagtttt attctagggg 480
 aaggaaacaa aggcggccga gtgccagctg cacgtttggt gggatttggg catcaggggt 540
 ggacatgctg cccaatggag ctgtcggcag ttgacccag cttggtcgt cgcgtcccga 600
 a 601

<210> 5
 <211> 801
 <212> DNA
 <213> Homo sapiens

<220>
 <221> unsure
 <222> position is 467 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 477 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 512 nucleotides
 <223> "n" refers to an undetermined base

<400> 5

gagtacgcgg ggcagaacca gcgcaatata gcattctggt aggggaacta attttgacta 60
 aaatatattgc caattctaata cccaattcc tggacctccg ggtagctggc aaggtatattt 120
 atgttagatg tgtctggagt aaggtgcacc ggagtatttc gacaagagac tcaattcaat 180
 gcgtattaaa acttgattga gagagggaga gagagaggtc attttataaa gaaagacctg 240
 tgaacactgt agattggaaa tttatgtttg caaaataaaa ggatgggttt atcaagtgga 300
 tgcatttaca aaatgtggca tccaggtttc gtaaaattag ctgaattcta cgggtaagat 360
 tatgaatgtg gtcataaat aattaatagg tagtgaaaaa gaatgtattt tgcattagga 420
 agtgcattca atagtatttc ggaaatgagc acttcgattt cctcggnttc catgcgnggc 480
 cacctctcca gagcagggca ggcacccagg gnggtgcccc cacaacaag cgcggtgtggg 540
 cattttcttg gctcgtgcgc tgaagtgcac gctgggcctt ggtgcccga cctcagcct 600
 gggagatagg gaggtggtgc tacctgcagg ccgattgtgt ccccgccata ggacactagt 660
 gggcggcaaa cctcacaaga ctcttcgagc cagccttcag cagagccagc aaaccagcc 720
 gccaccgagg gaggactgct ccatgcagat ggtcaggggc tttcttctga agacgcctcc 780
 cccacgatct ctcaagttca c 801

<210> 6
 <211> 204
 <212> DNA
 <213> Homo sapiens

<400> 6
 ccggttgctg ctggaggatg ggactacgaa ggatggggac tccgctcggc caccgctcct 60
 gaatggcctc taatctcggg gttaaataact ttatgagagt atcaatacca cctaatacctt 120
 tgctgagaat tactgctaga aatgtagatt ctgaggttcc gaaagtttgt ttttggttac 180
 cccctccagc tctcccgcg gcaa 204

<210> 7
 <211> 205
 <212> DNA
 <213> Homo sapiens

<400> 7
 agacctttat cgggcgtgag tgaaatcggg cgttcgtgtt tttcgtgggt cttccaacca 60
 caggccgct gagatggttc tagtcccttt gagtatacag acccttcctg tgcattgacc 120
 gacacagctc ggcccggatc ccgaaatgaa cgtttctacc ttcggaacgc tgcgtctcgg 180
 atccttctga acccgcacgt cgcaa 205

<210> 8
 <211> 597
 <212> DNA
 <213> Homo sapiens

<220>
 <221> unsure
 <222> position is 361 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 382 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 513 nucleotides
 <223> "n" refers to an undetermined base

<400> 8
 gaccatgaaa tcgtgtggct ctagccccct ctgggcctct tgttggtaat gaagccactc 60
 taaagcgccc cctgttattc agagggctcc ccagctgcca tgatatgtgt atggggaggg 120
 catagcaggt ctttttgecc cggcagccat tcttctgctc acaaggggct ggctctgggg 180
 acagggatgt ctttgtcatc agtgaccact aatccccctc ctcatgggcc tccagggctg 240
 ctcccccttca ctctcttggg tgaagttgta ggggctgagg ttaccctgag aaacacctgt 300
 tcttgagacc catagacca accttgagga tgcaggggga gccactggct gggctctgca 360
 ngtggggcca gctgatcccc anctgctggc acctccaggc atccacagag cttggagtc 420
 cagccacatt tctccttgg ccttagaggg agaggaagtc ctttgattgc ctagtccaag 480
 atccctttat ttctgccc gggattatgg ggnagcaagc catgcccttc atgggaagct 540
 gttctccctt cctcgggggt gggctctggc tcagctcggg caacagtcac gatgggc 597

<210> 9
 <211> 500
 <212> DNA
 <213> Homo sapiens

<220>
 <221> unsure
 <222> position is 10 nucleotides
 <223> "n" refers to an undetermined base

<400> 9

gccaaacgcn ataccctctg cggggtgaga atgcgggccc gcccggtcc tcccgtgagg 60
 ccagggcctc ctgttctcct agacacccca aggagccaac tcctcgcag aagttccccg 120
 cttctgctct tatttccaag cttcgcgctt tctacaaact ccctgttgcc ttgactttga 180
 tttccagccg tggtagagggt cagagtgaac cccggcgcg cccccgacgg catccccgca 240
 caccaggata ggagaaattg gagggcctgg ggctcgggc tccgcagtcg tcggaggaag 300
 aaccacccgc ggggtccgca agggaaagtg aagaggcccc ggatttttcc aaagcgctgg 360
 ccaggacccc gaaggaaggg gaggagtcac ctgaagccgg ggaaggcccc ttgggtgctc 420
 tgccttggat ccttatgttc actgactttc gcgaccctg gaggggggca aatccgcgct 480
 gtttccccca acttggttc 500

<210> 10
 <211> 343
 <212> DNA
 <213> Homo sapiens

<400> 10
 gccaaacccac accagtacct gggaccgggg ggagcccgtt cggccgcta aaccgggctg 60
 gctggcgcca gggtccggg aggtgcggtc cggcggggaa gccgtgatgg gaagcgactc 120
 tgtccaggga gtgtccttca ccaccacact cctcacgtcc aggcagtgat cgacggcctg 180
 gcggcaccct cacagcgggc ccatagcacg gggccacaca cgtcccctga gcttagcctg 240
 ggcacattcg tctgccgcg agggcttaag ccagtctgca gccgcgccc cgtcactcgg 300
 acgcaagtcc gtcgtccgct ctgccacgcg gccgctaagc cga 343

<210> 11
 <211> 291
 <212> DNA
 <213> Homo sapiens

<400> 11
 gtcctacaca ctccgcacac aacgcggccg gtgttaagtc tccaaacgcc ccgagagctc 60
 caaggaccgc gcgcgcgaag gcgccgtagc aagtgggcac acaccagaca ccaccccggc 120
 gtgttccgcg ggagaagcca gtgcacacat cctcccga ggcggggttg ccagtgaac 180
 acaggaatcc tgcccttttt ctagaaaagc cccctcccc actttccctc caatacactc 240
 acctgcgtct caacagtttc cttcttgogc tacacgcggc cgctaagccg a 291

<210> 12
 <211> 266

<212> DNA

<213> Homo sapiens

<400> 12

```

gtccggatca gtttccccgg ccaggtcgct tcccggctct aaccatttcg cgctctgctc      60
tgtccgctgg tttgtccctg cccggttcct ctccccgggc ctgtcagcct ccgctttctct      120
ggaggttcct gggactcatt tctgatccac cgtcttgctt tctctgggag catcgacttc      180
tctccatctt cgggctcact cctgactccc tcgctgccgc ccccgggggt ttccacgcgt      240
gtctctaacc gcggccgcta agccga                                           266

```

<210> 13

<211> 553

<212> DNA

<213> Homo sapiens

<220>

<221> unsure

<222> position is 497 nucleotides

<223> "n" refers to an undetermined base

<220>

<221> unsure

<222> position is 513 nucleotides

<223> "n" refers to an undetermined base

<220>

<221> unsure

<222> position is 517 nucleotides

<223> "n" refers to an undetermined base

<220>

<221> unsure

<222> position is 519 nucleotides

<223> "n" refers to an undetermined base

<220>

<221> unsure

<222> position is 527 nucleotides

<223> "n" refers to an undetermined base

<220>

<221> unsure

<222> position is 546 nucleotides

<223> "n" refers to an undetermined base

<400> 13

```

gatcctggtc catcgaaacc ttgtgtgcat cggttagtgc ttctggggcg tttgcttcta      60
gccgacgctg acagtggagt gccagaaaga gggagaggac cgtcatggct actctgcccc      120
tggtgtcacc atgcgtcttc ccccggcacc ggcgaggcga aacgttttcg tagtccccgg      180

```

gaggccccctc ggtcagggca gcagcatccc tgcaccctct ccgcagggtgg tctccccgac 240
gccacaggtg gccagcaggg cgcgggtggg ggcaggagcg cctctcccct gccaggcct 300
cccgctcctt ctcgagcgcg tgtggcgggg tggagagaca gccttctaca gctagtctag 360
ctcggcgcgg ttcccgtctg tggcctccta atcccacagc cacagcgctt tcctctaacc 420
tccctcggtg ggcttaaagc ctcccgttcc ttctgtctca ttcttctgc tccctcccc 480
cgaaaccccc agatganagc tgggaacctg gcnccantna ctgagcnaac agtggtgacg 540
ggccgnggcc caa 553

<210> 14
<211> 156
<212> DNA
<213> Homo sapiens

<400> 14
ggcacacag tgggtacaag gatgagctcg gtgtaaggaa tggaaagccc ccagtctaaa 60
ccaccgcccc ctagacacgg gtgaaaacct gcctaaaagc taactcaggc agtgactcta 120
tcaccgaag gggccctggg ccgcggccca agccga 156

<210> 15
<211> 300
<212> DNA
<213> Homo sapiens

<220>
<221> unsure
<222> position is 117 nucleotides
<223> "n" refers to an undetermined base

<220>
<221> unsure
<222> position is 154 nucleotides
<223> "n" refers to an undetermined base

<220>
<221> unsure
<222> position is 163 nucleotides
<223> "n" refers to an undetermined base

<400> 15
gttcacagcc cataaggtgg ggggtggccg aacctgaaac ggagcctgag ccaggatcct 60
gcaaccaaag tctgaagcgc ccccgggtgg gggccgagag cgctgcaggc aggtggnggc 120
gcgggggcagg cgggcgggcg aaggagctc cggntacgca ganaacgcgg agcgcctcct 180

tcccacctgc gcgagggcat cctgcccggg ggaggaaagg cgggagtccg aggcgggtcg 240
gattcccagc cagctccctc ctcacaggag gcgggccatt atccggcgtc gcaaagccga 300

<210> 16
<211> 196
<212> DNA
<213> Homo sapiens

<400> 16
ggcgcccagc aggggagcga gggaggaggg tgcagaaaga ggctccgaaa ttgggggaaa 60
ctgaccctg cttctctacc ttcggaggtg ggacagttgc acgaagtgt agttagaccg 120
gatcagttgg aactgacgga ggactgcaaa gaagaaacta aaatagacgt cgaaagcctg 180
tctcggcgt cgcaaa 196

<210> 17
<211> 299
<212> DNA
<213> Homo sapiens

<220>
<221> unsure
<222> position is 21 nucleotides
<223> "n" refers to an undetermined base

<400> 17
acaccaggag aggggaagaa nccagcacct accgacaggg gtggagctgg gtcaagaatg 60
gtgtgggtccc tgctttgggg gaatgctggg gaggtagaaa gcccttcta acggggcgtc 120
actgcaatta ctgcttctc tttcccataa aactccccct agtgtatcag aacccccaaag 180
gagtttcagt aagcggttct tctgttgtct ccggtgaga ctccagggga acctcaagct 240
cacatggccc tggccgggcc cctgggcagg agcaggcgag aggtctgcgc ggccgctaa 299

<210> 18
<211> 363
<212> DNA
<213> Homo sapiens

<400> 18
gggtatgtgt tacacatccg agataactac acaggcatcg accctgtcca cccggggatg 60
ctagaggggc tgcgctggtt ttactccagg ccatggtgag agccaccgtg aacacagggc 120
tctctcctct gagctgcaga agctctgtgc cctgtcccct gccacaagtc acagactttc 180
ttcatgtgtt ttacctcatg ttaatgaagg agatcttctc caggggcttg atctagtggg 240

aaacagagga ggggggggatt ttaaatttca gtccgtccaa ccctgtagat ctgctgtcct 300
 acagtaacgt aaaggatcac caggtaaaac gctgcttctc ccggacgccg ccccgcaagc 360
 cga 363

<210> 19
 <211> 322
 <212> DNA
 <213> Homo sapiens

<400> 19
 ccggcccgtc cctcttaata tggcctcagt tccgaaaacc acagaataga accgcggtcc 60
 tattccatta ttcctagctg aggtatccag gcggctcgga cctgctttga aactcctaata 120
 tttttcaaag taaacgcttc gggctgcagg aactcagct aagagcatca ggggggcgc 180
 aagaggcaag gggcggggat ggggtgggtggc tgcctcgtg gcagaccgcc cgcccgcctcc 240
 caagatccaa ctacgagctt tttaactgca gcaactttaa tatacgtat tggagctgga 300
 attaccgcgg ccgctaagcc ga 322

<210> 20
 <211> 255
 <212> DNA
 <213> Homo sapiens

<400> 20
 taataagata ccaaatacggg cgagaaacga aaagctcctg gcctccgtat ttggggccag 60
 agacaccgca gggagtcagg tccccgccga caaatcgga gaggcctgcg ggagttagcc 120
 agataatgct ctccctgtcc taccgctccc caccaatttg ccttttacct gccgcagagc 180
 ttgcttgaac caaaggggtt tgcggtcttc tcctcctcaa cttgcgatcc ccaggccttc 240
 gcgtcccgaa gccga 255

<210> 21
 <211> 406
 <212> DNA
 <213> Homo sapiens

<220>
 <221> unsure
 <222> position is 6 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 7 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 18 nucleotides
 <223> "n" refers to an undetermined base

<400> 21
 atgtgnaag gctcgetntc catttctctt ttctctcttc tccctctctc atgtgcggtc 60
 tccctcaaca tccaaaccaa ccgagtgcgt ctgaggtgaa atcgtgccag acttagagac 120
 ggctgccagg tttctctcaa gtcttggtt aacaaaagaa agcaaattac aaaaatggaa 180
 attttcaaac tagcgttcag tggatttcaa atcgacgttt gggtagcgca caggcacaga 240
 ccgcattcgt gctattttgt gattaaaatg ataccaaaaa tacctccttg ctttggtttt 300
 cgtcttcgaa aacgacttct ttcttcttc taatttcccc cttacttttg ggagcggcaa 360
 acccctgacc actctagaat tgctaacatt tggaccggcg tcgcaa 406

<210> 22
 <211> 210
 <212> DNA
 <213> Homo sapiens

<220>
 <221> unsure
 <222> position is 10 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 13 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 14 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 25 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 40 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 46 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 47 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 50 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 76 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 95 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 207 nucleotides
 <223> "n" refers to an undetermined base

<400> 22
 gcacgttcgn gcnnctgtga ccatnagctg ccaactggan gcaccnnggn aaggggtgggg 60
 gcctcctgga gacttngggg agagggatag ccggnataag ctctgtcct ttctataggc 120
 ataagcgggt ggtcaccacg gattggggat cccgaatccc tggctccaga tagacttaat 180
 gaagaagcac ctggatccgg gccgcgncaa 210

<210> 23
 <211> 310
 <212> DNA
 <213> Homo sapiens

<220>
 <221> unsure
 <222> position is 9 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 11 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 32 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 79 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 80 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 120 nucleotides
 <223> "n" refers to an undetermined base

<400> 23
 tcacgcttnc naaggctctg aatcctgagg gncagatctc caagaaggag ggaggctggt 60
 cctagtcccc gaggtcctnn actagggtcta gatcactggg taaaagaagg ggagcggcan 120
 cacgtatggg gtaggcgctc tcactactca catctcgaga cctttgccgg cgtagggctg 180
 tccgggggga acgacccgcc ttttccggta tcggttgta tggcggcgcc cagcccagcc 240
 tggttttttc cggtagccaa ttgaactaac aaccccgttc cctttaggac taatctgtca 300
 cgtcggcgca 310

<210> 24
 <211> 304
 <212> DNA
 <213> Homo sapiens

<220>
 <221> unsure
 <222> position is 13 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 74 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 266 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 269 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure

<222> position is 292 nucleotides
 <223> "n" refers to an undetermined base

<400> 24
 ctctggtctg tgntggatac gcgtgttctt ctgcggagtt aaagggtcgg ggacgggggt 60
 tctggactta ccanagcaat tccagccggt gggcggttgg cagtcactta aggaggtagg 120
 gaaagcagcg agcttcaccg ggcgggctac gatgagtagc atgacgggca gcagcagcag 180
 ccagcaaaag ccttcgcaaa gtgtccagct gctgcactgc cgcggggact cccacagcac 240
 catgactagt tcgtgcgact ctgcancanc aaacggcttc cgaggaacac angatcgagg 300
 gggca 304

<210> 25
 <211> 379
 <212> DNA
 <213> Homo sapiens

<220>
 <221> unsure
 <222> position is 6 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 10 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 13 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 19 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 21 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 31 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 113 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 184 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 206 nucleotides
 <223> "n" refers to an undetermined base

<400> 25
 aaaacncatn tgnagagcnc ntcggcagag ncgcagctgg ctgacccagg agaaggcgcg 60
 ctgggtgtgg ctgggacggc caaggccgcg gcttcccgcg tggggatgcg ctntggcgca 120
 aagctggtcc cggcggggcc aggcgtttgt gggcgggtga cggggatcta gggcttccgc 180
 tcgngattcc tcttgggctg tctttncggg tttggactcg cctgccaggc tgtgtgcagg 240
 gttcccgcgtg cctctggccg gcaggcgctc gggtgcagg tgggccggca ggcaggtgtt 300
 agcgggaagg gagcacaggt agcgaggtgg gatcggcgac ctggctaggg tgcggcaga 360
 atggaatgcg cggccgcta 379

<210> 26
 <211> 625
 <212> DNA
 <213> Homo sapiens

<220>
 <221> unsure
 <222> position is 8 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 18 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 50 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 64 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 609 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 616 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 618 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 621 nucleotides
 <223> "n" refers to an undetermined base

<400> 26
 gggacgcnag ccagggantt tgatccgttt tgaatgaaaa gaaagagaan ccaaaccaaa 60
 cctntcagtc atccaaaacc ttcaggcttc cagggagggtt ttgctataat tttctctaag 120
 catgactggt tctgggggag gggaaagggg tggttgtatt tactgaaaat tcaaatcgaa 180
 ataataaatg gccaaatttg gacacttacg gacccaaaca gttttgctca cgccagagaa 240
 accgagagca cagggtctgc gtgaagccta tctcggcaga aggcaacatt ctaataaagc 300
 ccgtgggaaa acagattaca ttttcgccat gaataagtca tgcagtgaaa aatattgcct 360
 acagcctgtc gacttatatt attatcacgt ttttcaactc ggcgtgagga gggagaggag 420
 tgttcatatt tgactaggaa ttgcaggatc gatgcaaact ccagggcagc agccagactg 480
 gcatatgtgg ggctctccgg ttactttctc tgtatgtcgc gggtagagag aacagcgagg 540
 acaatttagc gcaaacacac gaagggtcgg atctcaaggg ggcagcgctg ggagaaaggt 600
 tagggctgna gagcgnanag ncaaa 625

<210> 27
 <211> 499
 <212> DNA
 <213> Homo sapiens

<220>
 <221> unsure
 <222> position is 2 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 7 nucleotides
 <223> "n" refers to an undetermined base

```

<400> 27
gnctccnctg tcccctcggg cggaacggag gcaactttcc ggagtctatt tttgttaaga      60
caatcaactc caataactga gctgaagttt ttgtttaaaa agaaaaaaat ctgataagtg      120
atgattttac ctacttgtgg aactagatt tcaattagga aggttttttt aaacggcttt      180
ttgtaacttc gctgcaggaa gcaggtttgt ttctttttct tttcttttta agagaagggtg      240
tatttcactg gtgcaatggc ttggcacctc cggggcctgg gaggacctca gacctcccca      300
gccctgggtt tctcgtctt caagaccaac taggaagggt caagcgggga gagggagtgg      360
agggtcagggt gagatctcag agctgccccg gccggcccc gtctctttct acctctctt      420
ccagagaacc agcggctcac acctttctca acgcaggaca tgctcggcgg ccaaagccga      480
attctgcaga tatccatca                                          499

```

```

<210> 28
<211> 561
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> unsure
<222> position is 20 nucleotides
<223> "n" refers to an undetermined base

```

```

<220>
<221> unsure
<222> position is 21 nucleotides
<223> "n" refers to an undetermined base

```

```

<220>
<221> unsure
<222> position is 23 nucleotides
<223> "n" refers to an undetermined base

```

```

<220>
<221> unsure
<222> position is 26 nucleotides
<223> "n" refers to an undetermined base

```

```

<220>
<221> unsure
<222> position is 39 nucleotides
<223> "n" refers to an undetermined base

```

```

<220>
<221> unsure
<222> position is 40 nucleotides
<223> "n" refers to an undetermined base

```

```

<220>

```

<221> unsure
 <222> position is 44 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 49 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 65 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 80 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 98 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 107 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 471 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 484 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 544 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 559 nucleotides
 <223> "n" refers to an undetermined base

<400> 28
 gggcgattgt tattcaaacn ngntanctct ctgcgggggnn gagnaatgng ggcctcgcac 60
 ggctncatcc cgcgcgagcn cagggcctcc ctgttctnct agacatncca aggagccaac 120
 tcctccgcag aagttccccg cttctgctct tatttccaag cttcgcgctt tctacaaact 180

```

ccctgttgcc ttgactttga tttccagccg tggtaggggt cagagtgaac cccggcgcg 240
tccccgacgg catccccgca caccaggata ggagaaattg gagggcctgg gcctcggtc 300
ccgcagtcgt cggaggaaga acccaccgcg gggccccaa gggaaagtga agaggcccgg 360
gatttttcca aagcgctgcc aggaccccga aggaagggga ggagtcacct gaagccgggg 420
aagctccttg ggtgctctcc ttggatcctt atgttcaactg actttcgcg ngccccctgg 480
aggnggaaaa tccgcgctgt tcccccaac ttaacttcac gcggccgcta agccgaattc 540
tgcngaaatc attacactng c 561

```

```

<210> 29
<211> 717
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> unsure
<222> position is 643 nucleotides
<223> "n" refers to an undetermined base

```

```

<220>
<221> unsure
<222> position is 651 nucleotides
<223> "n" refers to an undetermined base

```

```

<400> 29
actctccgcg gtntcntggt gcctcacagg aggtgggggt cctccacccc ggtccccagg 60
cctctccctc tgcccagact tcccggctct gcctccttcg cctcgctgc ctgcccgact 120
ctgaaccctg ctctcttct aactaaaagt cagtgtttta tttctccgc agtccaatgc 180
ccgcgtttta cttattcaa taagaaggcg ttcatttatg gcaagacagg acagccaggt 240
aataagggcc tctgcacacg cgggccatt ggagggcg aactgcgaag tcttccgga 300
agagcttcct ggagagaagg ggaacgagcc agcgtttatt gagcatctat tatactaagc 360
atctgcttgg cagttcacga cggtcgcatt tttcatcct tacagcgatc cctatttgtt 420
cgcttgcttt aaagcctcac agctcacaaa gggctgggat ttattccaga tctctctctc 480
agatgccatc tcaattccag gtgtctctgc tgctttgaac gcgggaaacc cagcaaagg 540
agtgatttcc aaggccttct gtttgaata tctttaatcc tccccttatt aactggaaaa 600
actcccacgc atccttcagg gtcagctca aatgtccttt atntctgcag ngaaactttc 660
ccaaggaaaa ttagttacac agctaatttt agataaattg agccagttga tagaatt 717

```

<210> 30
 <211> 280
 <212> DNA
 <213> Homo sapiens

<220>
 <221> unsure
 <222> position is 30 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 189 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 192 nucleotides
 <223> "n" refers to an undetermined base

<400> 30
 tgatggatat ctgcagaatt cgggctttgn gacgccgggc acgcagtagg gaaaacagta 60
 ttaaaacgcc ctacagaaaa tctcggcgaa gtccccggag aactctgggt tctaagatca 120
 gctgggcgca ctttctccgg gacgtccctt cttctcggtc tcagcgcctt cctgccctca 180
 gccgcgceng tnttgttttg gtggcaaact gaaataagaa atggaaatat attggccttt 240
 gctgctgccca gggatgagag gttgttgacg tcggcgcaaa 280

<210> 31
 <211> 270
 <212> DNA
 <213> Homo sapiens

<220>
 <221> unsure
 <222> position is 2 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 5 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 6 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 8 nucleotides
 <223> "n" refers to an undetermined base

<220>
<221> unsure
<222> position is 9 nucleotides
<223> "n" refers to an undetermined base

<220>
<221> unsure
<222> position is 11 nucleotides
<223> "n" refers to an undetermined base

<220>
<221> unsure
<222> position is 12 nucleotides
<223> "n" refers to an undetermined base

<220>
<221> unsure
<222> position is 24 nucleotides
<223> "n" refers to an undetermined base

<220>
<221> unsure
<222> position is 26 nucleotides
<223> "n" refers to an undetermined base

<220>
<221> unsure
<222> position is 27 nucleotides
<223> "n" refers to an undetermined base

<220>
<221> unsure
<222> position is 29 nucleotides
<223> "n" refers to an undetermined base

<220>
<221> unsure
<222> position is 33 nucleotides
<223> "n" refers to an undetermined base

<220>
<221> unsure
<222> position is 36 nucleotides
<223> "n" refers to an undetermined base

<220>
<221> unsure
<222> position is 227 nucleotides
<223> "n" refers to an undetermined base

<220>
<221> unsure
<222> position is 244 nucleotides
<223> "n" refers to an undetermined base

<220>

<221> unsure
 <222> position is 245 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 264 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 265 nucleotides
 <223> "n" refers to an undetermined base

<400> 31
 gnggngngna nncggcgatg gatntnngna ganttnngtg atggatatct gcagaattcg 60
 gcttagcggc cgcgaacaaa gagcgaacca aaggatgctt cgaattttta aaacggaatc 120
 tctgcacca aatgcaggac tggtgactta aggagctgcg aagtctgatt taccgggcct 180
 actctcgacc tgccccccac cccagctca gggggacctt tttatcntga acgccagagc 240
 tacnnaccaa gtcgggtggc cacnnccaaa 270

<210> 32
 <211> 347
 <212> DNA
 <213> Homo sapiens

<220>
 <221> unsure
 <222> position is 7 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 8 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 11 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 50 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 309 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 313 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 322 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 325 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 331 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 336 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 337 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 338 nucleotides
 <223> "n" refers to an undetermined base

<400> 32
 tttggannta nngggggcgtg gcgtggatcc agtttccccc ggccaggctn gcttcccggg 60
 ctcaaccatt tgcgctctg ctctgtccgc tggtttgtcc ctgcccgggt cctctccccg 120
 ggctgtcag cctccgcttc tctggaggtt cctgggactc atctctgac caccgtcttg 180
 cgttctctgg gcgcacgac ttctctccat cttcgggctc actcctgact ccctcgctgc 240
 cgccccgggg gtttccacgc gtgtctctaa ccgcgggcgc taagccgaat tctgcagata 300
 tccatcacng aantctgcag anatncatcg ncgaannnca ccgcact 347

<210> 33
 <211> 342
 <212> DNA
 <213> Homo sapiens

<220>
 <221> unsure
 <222> position is 193 nucleotides

<223> "n" refers to an undetermined base

<220>

<221> unsure

<222> position is 299 nucleotides

<223> "n" refers to an undetermined base

<220>

<221> unsure

<222> position is 300 nucleotides

<223> "n" refers to an undetermined base

<220>

<221> unsure

<222> position is 301 nucleotides

<223> "n" refers to an undetermined base

<220>

<221> unsure

<222> position is 302 nucleotides

<223> "n" refers to an undetermined base

<220>

<221> unsure

<222> position is 325 nucleotides

<223> "n" refers to an undetermined base

<220>

<221> unsure

<222> position is 328 nucleotides

<223> "n" refers to an undetermined base

<220>

<221> unsure

<222> position is 337 nucleotides

<223> "n" refers to an undetermined base

<220>

<221> unsure

<222> position is 338 nucleotides

<223> "n" refers to an undetermined base

<220>

<221> unsure

<222> position is 339 nucleotides

<223> "n" refers to an undetermined base

<400> 33

gtagggcgcc gccgtgacag attagtccta aagggaaacgg ggttgtagt tcaattggct 60

accggaaaaa accaggctgg gctggggcgcc cgccatgaca accgataccg gaaaaggcgg 120

gtcggtcccc ccggacagcc ctacgccggc aaaggtctcg agatgtgagt agtgagagcg 180

cctaccccat acngtcggcc ggctcccctt cttttaccca gtgatctaga cctagtctag 240

gacctcgga actaggacca gcctccctcc ttcttgaga tctgaccctc aggattcann 300

nnctttgctc acgagctcca acccnacnca tccaaannnc aa

342

<210> 34
 <211> 370
 <212> DNA
 <213> Homo sapiens

<220>
 <221> unsure
 <222> position is 325 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 343 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 361 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 368 nucleotides
 <223> "n" refers to an undetermined base

<400> 34
 cattgtttac tttcgtctaa acgcggtgga agcccatgga agaaagcggg tagcagcaag 60
 gcagagccct gctccctctg cagccccagc tccagcgcc ctgggctttc caggcacctg 120
 tccgggtagg ggattgaggg ccgtggccag gccgcactt tctgctagc cgcagctggc 180
 cacatgccca tctgaccctc cgagttctcc tctaaaaatg gggctgacag ccgctacctc 240
 acaaagtcca caccgggctc aacccgntgc ctctctccc aacaggactc tgccaccctc 300
 cctcaggatg cctgagggcc ccganctgca cctggccagc cantttgtga atgaggcctg 360
 nggggcgntt 370

<210> 35
 <211> 213
 <212> DNA
 <213> Homo sapiens

<220>
 <221> unsure
 <222> position is 8 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure

<222> position is 10 nucleotides
 <223> "n" refers to an undetermined base

<400> 35
 aaaatacnan taaagcgatg cttcgaattt ttaaaacgga atctctgcac ccaaattgcag 60
 gactggtgac ttaaggagct gcgaagtctg atttaccggc ctactctcga cctgcccccc 120
 acccccagct caggggacct tttgtctgaa cgccagagct actgaccagg tcggggggcc 180
 gcggcccaag ccgaattctg cagatatcca tca 213

<210> 36
 <211> 173
 <212> DNA
 <213> Homo sapiens

<220>
 <221> unsure
 <222> position is 4 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 5 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 100 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 109 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 123 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 144 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 156 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 160 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 162 nucleotides
 <223> "n" refers to an undetermined base

<400> 36
 gacnncgggt ttgtgtgtaa cagggtcagt ccccgatatct actttgcgaa agcttcgagg 60
 cgagcgtgaa gtcaagggct gcggtggatg ggggtaaaan gcctcctcnt cccactgcct 120
 gcnccgtctt ggggtaaccc ctanccccc cccgngnttn cnctttaatg etc 173

<210> 37
 <211> 369
 <212> DNA
 <213> Homo sapiens

<220>
 <221> unsure
 <222> position is 22 nucleotides
 <223> "n" refers to an undetermined base

<400> 37
 tcactgtgcc ggggtctctcc tncccggtcc aactccetta cttgtcctca tctctgtccc 60
 caagggtccgt gacccgcgga ggtgatgggg gggataggag agccccaggg accgcagagg 120
 tgacacaatc gcccgccgt cctccctcgc tgggagccga ttcagcctgt gccgagcctc 180
 tcccttcgcg tgctctgcg cacagcgggtg gcaccgcagg actccgggtc cccccggct 240
 ctccatcggg aagccggcaa atgcgcttcc tcagccagac cgcggcgggg tgggggcggg 300
 gggggcgga gttgaaatac tgggacagaa acacctgcc gtcccaaggg acggaaaact 360
 ggatgcaa 369

<210> 38
 <211> 123
 <212> DNA
 <213> Homo sapiens

<220>
 <221> unsure
 <222> position is 20 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 29 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure

<222> position is 41 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 87 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 108 nucleotides
 <223> "n" refers to an undetermined base

<400> 38
 gtccttcgc cccgcttttn ctttcccnna ggtcccagcg nccgaaccgg cggatgtcca 60
 cgaaacatag ggcgagccgg gggccangcg gggccgtgta aaatctcntg tggtcatttt 120
 gtg 123

<210> 39
 <211> 450
 <212> DNA
 <213> Homo sapiens

<220>
 <221> unsure
 <222> position is 32 nucleotides
 <223> "n" refers to an undetermined base

<400> 39
 ctagccctgg aagagaatcc gaggtcagc cntgctgcag caccaggac actgcatccc 60
 agcacctgcc cgaagatcag cccaggaccc aaaggaaagc aggctccaag ctccccggaa 120
 gccaaggaaa ataggaaaac atatcctgcc ccggggacac cttctggaac tatgaccaca 180
 tgcacttgac cttccggaac aatcaccgca tgcacctgac ctcccggaaac tgtcaccacc 240
 gcgcgcacct gacctcccg cactgtcacg accgcgcgca cctgacctcc cggcactgtc 300
 atcaccgcgc gcacctcacc tcccggaaact gtcaccaccg cgcgcacctg acctcccggc 360
 actgtcacga ccgcgcgcac ctgacctccc ggaactgtca tcaccaggcg cacctgaccc 420
 cccggcactg tcacgaccgc gcgcacctca 450

<210> 40
 <211> 593.
 <212> DNA
 <213> Homo sapiens

<400> 40
 ggaccaagct gggtaaactg ccgacagctc cattgggcag catgtccacc cctgatgacc 60

aaatcccacc aaacgtgcag ctggcactcg gccgcctttg tttccttccc ctagaataaa 120
 actccgctgc tttccacgt tcctggagca gcagccggaa taaagcgccc atggccttgc 180
 cctttgagtc tcggaggatg tttgccactc caacaatgga cttttaaata attcaggggt 240
 caaaaggcgt gtgtgtgggg ggggagaaaa gttacaaatc agcacttgaa accgaacaca 300
 aacaaaaatc aaacaaatcc gaactaatat aacaaatcaa aactttgatc tttagaagaa 360
 aacttcaacc ttaatgcttc caggaggaaa gcagaaagga taatgactga attgtgaaaa 420
 cgagccaaaa tgttccacca ctgatgtcac acacacctat gactccctgc acagatccac 480
 ggtcccgggc gctgaatccc cgcaaccctc tgcgccaca gaggttaaac tctcgctgct 540
 ggcgacttcc gcttctggc ctaaactctga cacgcacgac tcccccgcg gca 593

<210> 41
 <211> 457
 <212> DNA
 <213> Homo sapiens

<400> 41
 accaccaacc aaatagggcc tttcctgtta acgaccacgc ggcaaggggg ccgggcccctc 60
 gaacgcctcg acggcctccc ccactccaaa gggactccga tttcgagga tctcccgctt 120
 cccgcctctg ctcccaacac cctacgtttt tctcttctc ctcatttacg tatttacaat 180
 aaaacagcga agctgcacag tctgtctcta aatcaaagc gggtaccatc aaagcctcag 240
 actctatgtc tcaaccgcaa aaggctctgac aggaaatcaa ctcgggagtt tgtcaattct 300
 ttaaactcaa agctctgtta acgaaatctg gatctttcct cgctccccac ctgcctcccc 360
 tgacaggaga atgactgtaa aaggatcctg tcgtccccga aagtcagcac caagcacttc 420
 acaaattgtc aaatctcaaa agcttacacg cgcggca 457

<210> 42
 <211> 211
 <212> DNA
 <213> Homo sapiens

<400> 42
 gcctgacctg aatgacgcgc atgttgaggc cggctctctg cgccagctgc tcgcggatgt 60
 ggcgggtggg cttgggtgta gcagcgaagg cggccttcag cgtctccagc tgcttggctt 120
 tgatggtggt gcgcgggtccc cgccgcttgg cgcccagggt ctggctgtca ttctcgttgc 180
 taccgcttc cttgtccgac acgtcggcgc a 211

<210> 43
 <211> 141
 <212> DNA
 <213> Homo sapiens

<400> 43
 aaatcatctc cgggggcccc gcacggacac gctccagacc cgtgagttcc ccagcgccgt 60
 gccgggaggt caggggcgct gaaagaagga agaattcagc cacctctcag catccctgtt 120
 acctcgagga cgcgcctctc a 141

<210> 44
 <211> 559
 <212> DNA
 <213> Homo sapiens

<400> 44
 acccactttc cattaacact aaataaaacg catccatgga tttcctctcc attccgagggc 60
 aacaggagtg catggcacat tgcctactc ccctgaagct cttogetaac ctaagactcc 120
 aggggtgagga agtttagctgg agctttttta agtgcatctc caaagagaat tttgctcaca 180
 ccatgagagc cccaagaaa caccagggcc cccttagatg ccggagacca cgccctccag 240
 gaataagccg caccctctgc ccagcagatc cttgcgcgag tagccctctt tccctggggc 300
 taatcaagtg catgccacat gtcaccactc tcagctggca attcttctc agaggcgag 360
 actttcacgg aatccccagc aggggggggtt aagagattca ggggaggccc cgcccggtgcc 420
 ttccacaaaa gtgcgtttac cgtggctcgt gtcctgcggc cccaaggggg tagcctggga 480
 cgtgtattgg gagggcatag aggtccttc caggacaagc tgccagcctc cagtgggcaa 540
 ccatgtgaga ggcaaaatt 559

<210> 45
 <211> 433
 <212> DNA
 <213> Homo sapiens

<400> 45
 gcgaacagca caaaggcttc attcctacga gagattaagt tttagagcaa atggacacga 60
 tcgttaaaga atttgatatt tccatgtaaa ctgcattagc aggttatgag atccaaactc 120
 acaggaacaa ctccaactct cgcccatgcc ctatttcatg tctagatttg ttttaaccgac 180
 ttacatcata atccaagaat acgaactaca gtatattctt acagcaaagt tattccttaa 240
 aagcaaaacc gagccacctt tgaaaacacg cacacacatt atccacggca ctaaaacccc 300
 agtcttgacc gagaaagacc aacaacttgg gggggaagaa aacaacttca gagccagagc 360

tcccaaagca gaaagcgctg gcggctgaag ggcacacgag gttccgctcc cgggcgaacg 420
ggcggcgctcg caa 433

<210> 46
<211> 487
<212> DNA
<213> Homo sapiens

<400> 46
cccttagtat tccatgagcc accatthttcc ccacgatccc tccagcctga acgatcacat 60
cctactgtgg accacgactc tcccagcagc gggcgthta tatecagtha gcaggthtctc 120
accacccccct cgctggctcg aatacagcat ctgcaccgag ttcccagaga tegtcaaccc 180
agcaaattccc ttaattgggtg gacatgaaaa tccagggctt tgtgctgtaa taacagagtc 240
ctggggggcct ggggagthttg tgccgcttg agctcaggth tctgggacag aggctgagcg 300
cagggcaggg aggcaggtht cacctggcac ctcccagag cctcgccgag cagatggaag 360
cagaggctct cgcgcccggc ccccgccggg agacctctct ctctttccct cggcctgctc 420
tgccctctcc cgccttctcc ctgtctgatc cttctctgct gtcattgtct ttgtcctcgc 480
gccccga 487

<210> 47
<211> 403
<212> DNA
<213> Homo sapiens

<400> 47
gtcatataag cacaaccatt cccagggcca ccctggatgc atcagatcag tccccccact 60
ggtgaccaca atggctggct cagagtgcct ttgaacagac aggagaaaca gacttcttgg 120
agggagggac cttcccacag ggaatggcca aggagctagg tcttcagggc ttgcatggcg 180
tggagtgtgt gctcaggthc acagtgaagc aaacctgagg ggacttgggc cctgcgtcct 240
ccagcacaca cgcacccttt cgcgctcaca tccggggcac ccacccgtgg aatatgtgag 300
ccgcacttgg ccagccaaga gttccagggc caggaagthc tgcttctcgt tcaggcgccc 360
gttgtagaag agcagcccgc tctgctgcac tgtcgcgtcc cga 403

<210> 48
<211> 155
<212> DNA
<213> Homo sapiens

<400> 48
 ggcggtggaga ggaggggggca gaaactcagc cgccccctacg tttgctaaac tgcgtccgcc 60
 agggggcgta tttttctaaa acgcacaaga cgtttcgtgg gttatcgatg gtctcttgag 120
 cctccttgac tgatggggat tgaccgggag ggata 155

<210> 49
 <211> 256
 <212> DNA
 <213> Homo sapiens

<400> 49
 tctactgagc ttttctttaa gtggaaccag aagtgcgtgg atgagaggga aaggatggga 60
 gtgcgtccaa aggtggacag caggtcccca tccctgggtg gagtgagact ggacggcacc 120
 ccccggaag gtggtttggg ccttggacaa ggctagaggc aggagtccat gatgcagaga 180
 tgacacagtg cccctccgag tgtgagtgca cgaagggtcac tactgaggct ttgtgcttgt 240
 aaaaggccgc ccgca 256

<210> 50
 <211> 224
 <212> DNA
 <213> Homo sapiens

<400> 50
 tgcgggggtcg tgggggaacc ggcgggagct gtgcgtggc cggcctcact ggagtaggaa 60
 ttttagatga aactgagtc gtttctcctt gaaggcaggc agtattctta gatctactat 120
 tcatttaaaa agaaggaaaa gaaaaaaaaa tgactgctac ttactgagaa gaaaatttct 180
 gttctcctcc gattccgctg atcccgcttt atccgcgcac ctca 224

<210> 51
 <211> 313
 <212> DNA
 <213> Homo sapiens

<400> 51
 gtggctggga cggcccaggc cgcggttcc cgcggtggga tgcgtgtgg cgcagagctg 60
 gtcccggcgg ggccaggcgt ttgtgggagg gtgacgggga tctagggtt ccgctcgtga 120
 ttctcttgg gctgtctttc cgggtttgga ctgcctgcc cggctgtgtg cagggttccc 180
 gctgcctctg gccggcaggc gtccgggctg caggtgggccc ggcaggcagg tgtagcggg 240
 aaggagcac aggtagcgag gtgggatcgg cgacctggct aggtgtcgg cagaatggaa 300
 tgccggccg cta 313

<210> 52
 <211> 385
 <212> DNA
 <213> Homo sapiens

<400> 52
 tacgttgccg attcattctg ccgacaccct agccggtcgc cgatgccacc tcgctacctg 60
 tgctcccttc ccgctaacac ctgcctgccg gccacactgc agcccggaag cctgccggcc 120
 agaggcagcg ggaaccctgc acacagcccg gcaggcgagt ccaaaccggg aaagacagcc 180
 caagaggaat cagcagcgga agccctagat ccccgtcacc cgcccacaaa cgcctggccc 240
 cgccgggacc agctctgcgc cacagcgcat cccacgcgg gaagccgcgg cctggggcgt 300
 cccagccaca cccagcgcgc cttctccagg gtcagccagc tgcggctctg ccgaagcgct 360
 cctccgctcc tttctcgcgc ccga 385

<210> 53
 <211> 307
 <212> DNA
 <213> Homo sapiens

<400> 53
 aaccgcgctc gggtcggcaa gggtcaggga gacaaggtag agaaggctgg ggtgagcaag 60
 aagtcggggc gccgatcgtc agggccacga gcctcgcctt gccttcttgg aatcccaccc 120
 aactttaaag gcccaaagat cctgaaaatt ccgaaagcga aactgcgggc tggctctccag 180
 aagtttgaga acggtctccc aggttttcca gcgtcgtccc gggattctcg gacaccacaa 240
 acgccatcaa ccacgagcac cgggtgtccgt ggctattgcc ccgaatggtc cccatccgcg 300
 tccccta 307

<210> 54
 <211> 182
 <212> DNA
 <213> Homo sapiens

<400> 54
 cgatgtcgaa gccgtttgga gggaacagcg gtttccaagt tcctgctgac ttgagaagcc 60
 tctgcgggtt tccgaatctc cggcgcactc ctgggcgcgc tgcgggagct gtagctcagc 120
 cagccaggga gtagcggctt tcatccgccg ggaggagtct ttcgagttca atcgcggggg 180
 ca 182

<210> 55
 <211> 523
 <212> DNA
 <213> Homo sapiens

<400> 55
 tcgggtttga tccgccccaa ccaaataagg cctttcctgt taacgaccac gcggcaaggg 60
 ggccggggccc tcgcacgcct cgacggcctc cccactcca aagggactcc gatttcgcag 120
 gatctccgc ctccgcctc tgctcccaac accctacgtt tttctcttcc tctcattta 180
 cgtatttaca ataaaacagc gaagctgcac agtctgtctc taaatcaaac gcggttacca 240
 tcaaagcctc agactctatg tctcaaccgc aaaaggctctg acaggaaatc aactcgggag 300
 tttgtcaatt ctttaaactc aaagctctgt taacgaaatc tggatccttc ctgcctcccc 360
 acctgcctcc cctgacagga gaatgactgt aaaaggatcc tgcgtcccc gaaagtcagc 420
 accaagcact tcacaaattg tcaaattctc aaagcttaca cgcgcgggca ctccggaaag 480
 gctgtgggga ccacccaaag caccctcctc cacaccgcgg gca 523

<210> 56
 <211> 795
 <212> DNA
 <213> Homo sapiens

<220>
 <221> unsure
 <222> position is 741 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 762 nucleotides
 <223> "n" refers to an undetermined base

<400> 56
 ttctactttct tccggctgac gtccatctcc tcaaatttct caggaatgtg gggaagctcc 60
 tagccctgcc tgcctttcta gagggcttct tggatttgca gcttctaaca agttgctctc 120
 gccacggaga agctgttatt atgacaaaat atttggggca ttatcaaaat cacacaggct 180
 gctgggctgc tgcgggtttc tcgccagggc cagtaagcag ttacatttgg agttgctacg 240
 tgttgttttg gggccgggct gtggagagtg actgagccag tatttttcat ccaaatttct 300
 gcaaattgaa ttaaccacaa ttctagtctc acctcccgtc tttaaaaaaa taagttgaag 360
 aaaaggtaaa tattagagat aaggcagcat ctagtgactg cggagaggca caagctggtg 420
 ggcgagggtt gggggagtca gcaaagccct taaaacctc cccgtttaat tttctggtg 480

tctctgcatc ctgttgccag aattccaaat gcttggagtc atttanaggt gcgagaactc 540
 aaacgtcgtt ccacttggaag aggggaccgt ttaacgttaa attccattag cacctaaatt 600
 gtttcttaaa gacatccgct cagacacagg actcgaaagc gagcatttca tgcaaataaa 660
 tttctcaaat tttaaacatt gttaaaagct tgtctcgacac ctcggtcccc tccccctccc 720
 cggaaganaa caataggccg ntggcgcatc cccacttcgg antaaatatt gacgggggaa 780
 gttgctaaaa acatc 795

<210> 57
 <211> 438
 <212> DNA
 <213> Homo sapiens

<400> 57
 gcctgtgtgt aggggactgg aggtggggga acctgttctt ttcttgtgtc tgatcctggg 60
 gctcgttcc tgggtcctag aacagcagcc aggacggaag aaactgttca cgttgcaccc 120
 ctttctctaa gattcccagg ccaagagtag ctgcagaagg tggccctgaa tctatggcct 180
 ctttctctct gcctgaccgg gctagtggat ccggagaggg gaccagggag agctcctccg 240
 agcaggggtc cttcgggaga cagagagggg tccaggctga gagaactctt caagcatggc 300
 gagtctgcgt tatagaatcg ggcgggcggc tcaacttggg ggaagcacca agaagagctg 360
 ggcgacctgg agcgcagaac cggctttggg gagccaccgg gcggggcagg ggtagcacgg 420
 agcccggggc gcggccca 438

<210> 58
 <211> 611
 <212> DNA
 <213> Homo sapiens

<400> 58
 gcttccccct tcctttctcc cgcgctgcc ccttgagatc cgacccttcg tgtgtttggg 60
 ctaaattgtc ctgctgttcc ctctcaccgg cgacatacag agaaagtaac cggagagccc 120
 tacatatgcc agtctggctg ctgccctgga gtttgcacg atcctgcaat tcctagtcaa 180
 atatgaacac tcctctccct cctcacgccc agttgaaaaa cgtgataata atataagtcg 240
 acaggctgta ggcaatattt ttcactgcat gacttattca tggcgaaaat gtaaactgtt 300
 tccccacggg ctttattaga atgttgccct ctgccagat aggttcacg caagccctgt 360
 gctctcagtt tctctggcgt gagcaaaact gtttgggtcc ataagtgtcc acatttggcc 420

atttattatt tcgatttgaa ttttcagtaa atacaaccac ccctttcccc tccccagaa 480
 acagtcatgc ttagagaaaa ttatagcaaa acctccctgg aagcctgaag gttttggatg 540
 actgagaggt ttggtttggt ttctctttct tttcattcaa aacggatcaa actccctggc 600
 tcgcgtcccc a 611

<210> 59
 <211> 291
 <212> DNA
 <213> Homo sapiens

<400> 59
 gagtttggca ggccccggat tccacaaagg agtaggcgcg gccagccgcc tccagccctg 60
 agctcagtaa attcgggtgtc ctgaatgctc ccttcctgtc cttaccactg cgagctctct 120
 tgggacagct ttctaggttc cactgcgacc tactttccgc tccctgagtg cttctttgct 180
 gaaactgcag gcgaaaagat ctctttccca gaccgcagcg cactttgaga aggggctcaa 240
 agtcgcccgc tctgaatccg gcaccggcaa ataggagtag ccgcatgcgc a 291

<210> 60
 <211> 226
 <212> DNA
 <213> Homo sapiens

<400> 60
 gaaaacagat aaaacgccct acagaaaatc tcggcgaagt cccggaggac tctggtttct 60
 aagatcagct gggcgcactt tctccgggac gtcccttctt ctcggctctca gcgccttcc 120
 gccctcagcc gcgcgcagct ttgttttggt gacaaactga aataagaaat ggaaatatat 180
 tggcctttgc tgctgccagg gatgagaggt tgttgacgtc ggcgca 226

<210> 61
 <211> 580
 <212> DNA
 <213> Homo sapiens

<400> 61
 ctgtgatgca ctccggcgat ctccgtggca gctgcctcct tcctctccag tgacgcctgc 60
 atgtgtcct aggcagtgtg aggagtgaag atgagatttg gcgcattctt caacggagtc 120
 tgagcaaage taaagggctc cgattcgtgc aagccaaggg ctgcccctcc tatcctgtcc 180
 tccttgagga cctgtgctaa ggcttttctca tccaccaggc caccatgggc tgcggttcaca 240
 aggaatgctc cctgtctcat ctgctttata gtaaagtcac tgacgaggtg gtgggttatgt 300

tcattgagat tgctgtgcaa cgagacacag tcaactctgat acagcaaacc ctgcagggtg 360
 tatcagggtc ccctctgcat gccctgggac ctctctatct tgtcctacaa gtaggggtca 420
 taaaatacga cgctgaatcc aaaggccttg gctcaaactg caaccgcctg cctcatgcaa 480
 ccgaagccca tgaggcctag cgtcttccac gaatgagggc cactcccatg gccacctcga 540
 gaatctgctc cacgctctga acccgcgcac ctcaagccga 580

<210> 62
 <211> 633
 <212> DNA
 <213> Homo sapiens

<400> 62
 gcccaggaga agccctccac ggtgggcgtc ctctagaca accagcacc cctgcaggca 60
 ccctcgtctg gcagaatcag ccctttccca cctgcaggcc cttctcagcg cctctgactt 120
 cccacacaca gcacagggtta caaactggtc cctggcagtg cactctagcg ggcctctctc 180
 acaagttctg cgggcctcgt ttcattgaaa gggggttggt gattcctgct gcccttggat 240
 ggccctgctg cagcacacc tctgagcggg cactgagcga gcgtggggag ctgctccctg 300
 ggaactaggc aggagctttt aaacaccctt acacacagcc attctgcggg aatacatgct 360
 ttcccggtaa ggcttttact gttcattcca ggtaaattgg aagtcgcaca cccaagctc 420
 caaatacaac tcgttagctg gcaggctctt gaagccaatt ccttctgagg aaaatggaga 480
 taatagcagc taccctccca ggtgactggg ggagaataaa gtggctgtgc atagtgggtg 540
 ttgcagctgg tggctgctat tctcttcat tacagcttgt aaaaaggggtg tctaggccat 600
 ttacacacag ataggccggg tggggtaagc cga 633

<210> 63
 <211> 703
 <212> DNA
 <213> Homo sapiens

<400> 63
 gcctatgaat ggatttataa ttgctttatt tttgtcccat ttagacagaa gtcagagaca 60
 gaggagagaa ccaaaaaact tggatgtttc cgtaaactag attcgtcaat cctcgataat 120
 tgaaagtagt tccagtatgt cagccaccgg ggttccctgg ggagctaacc agtcctgaag 180
 gaagtatgaa gaggaagagg aggtcttcag ttaaggggat gaatttgtgc agtcctaagc 240
 cctgcaaagg tgctggaggg aggaagaagg gcaggaaata aaagatggaa gaaaatttgt 300
 tttttatcca cttagagttt tatctttaat gatgggaaac agtgctgctc tcaggaaact 360

cagtgtggag atctaggagt tcacggttca tagtccatta ggagcaggaa aaggatagag 420
 gacatttata aagtaacatc caagtccaaa gtaaaatggt ataaattggt tcccatgata 480
 aaggctggct gagtaggtca ggaaaggctc tgtcagacca tatgtgctgt ttcaggctgc 540
 ttcaaattct tttaggacag tgggtggatat gagtgaagac ggggcaggca ggccacatct 600
 cttagaagag gaagggtgatt gccacgtctc ctctctccat gctgatggca aggcgtgcgg 660
 gctgtgttct cttgcagcca gegtcccatg ctgcggcgcc aaa 703

<210> 64
 <211> 420
 <212> DNA
 <213> Homo sapiens

<400> 64
 gtgacgtgcg gaatacacgt gatgtcgggg acaggagcgg gctgaagagg gcacgatccc 60
 acgcggaggg caccctcac ccggggtagg agcccgtgc acttgctgtc gctcagcccg 120
 ggcgtgcac cagggcagcc gccatgctgc ttaaagccgg tcatgtgacg cgggagccag 180
 ggtggaaggg gtccccgcgg gcaagccttc gacacgtgac ctgccaccg actacggaag 240
 cctcttgggc gttccgcccg ggtcacatg tcatgtgacg gccggccggt cgcggagta 300
 accaggaact ttcccagacc ctgcgggtccc tggagcgtca aaaagagcgt ccccgtagt 360
 aggtggagtc gcctgccctt ccgaatctca gctgtcttat ctggaacccc cagcgggcaa 420

<210> 65
 <211> 496
 <212> DNA
 <213> Homo sapiens

<400> 65
 gcgctgcacc aatttagagg gtagaaaaag gagttagaag caaagaggaa aaaataaata 60
 aacaggcaac aaaaacccaa ccagccagc ctgagccatt tgcattagtg ttcatttagg 120
 aaattagcag acgggaaacg ctggggagtg gagtgggccc cggccttggg gactgcagag 180
 cccgctcagc cctgggtggc tgggcccaca tggctgtgc caggagcaca ggaggaccca 240
 gaggtggccg agggagcctc gccgggctcc ggtatgggtc ctggcccctc acaggtgcga 300
 gcctggccca gtgactgtgg acgctgtggg agagcaggcc tccgatacgc agggctggga 360
 ctgctgacct ggaaggtggg gccgggcgtg tctggtgaag gcgcggttg cagctagaga 420
 gagacggcgg atggggtgac gccataaccc acggtcccag ttttgaggct tgacggtgac 480

ggaaaaggac gtcggc

496

<210> 66
 <211> 637
 <212> DNA
 <213> Homo sapiens

<220>
 <221> unsure
 <222> position is 612 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 627 nucleotides
 <223> "n" refers to an undetermined base

<400> 66
 cgccgagccg ggatgagcaa ggcttcctgg aggagagggc cggcctgagc ttggaaggat 60
 ggggaggagc cactggctac aaggggtgtag aggtgagaac cagtgtgacc tgcccatcgc 120
 tggctcgtctc tgggtcattc agctgaaatg gcatctctga gctgagagga gtgttgacctg 180
 taaggagcta ggcatcagcc ccagtagag gggcgggccca ggcacagccc atagccgcag 240
 acttagtgag tctagctagg gagacagtag aggggccaaa atgaggacac aggtcaccaa 300
 aaatcctggc caggtcctgc cactacctgg ctacgcgacc tgcccccccg agcctcagtt 360
 tccccattg gtggaatgga gtgaggaaga cgcgctccc ggggctgcga tggagaattg 420
 agtcagagtc tgggggtgct gggagggctg gggagcagcc tccctgagcc tcagtttccc 480
 tggctgggga atgaggacct tgctcgtccc cctcataag gggagagctgt caggaaagtg 540
 ctttcaacgc tgagccattt ccagtggtg cacaattagc tttccagagg attttggtgg 600
 attetagagc tngagggtg ggggatnggc ggccaaa 637

<210> 67
 <211> 595
 <212> DNA
 <213> Homo sapiens

<400> 67
 gccctgagct cttgagggcc tctgcagttc ttgggacaat tctgggacta tatctttggg 60
 ccttggtgag atctagaggc tctaaagtct ttgggagggg tcctgagctc cgtggacggc 120
 agggctcttg gcactcactt gcattcttga ggggtgtgtt tggcctcgtc cgtgcaggtg 180
 tagaatttcc cctgtagaga ggatgtctgt caagtaggtt cacccttcat cacactcccg 240
 ccagacccc tgcctggcat tccctccagt gtttgcccca ccttgaagag ctgcaccccg 300

atgcaggcga acataaattg cagaagtgtg gtgacaatca tgatgtttcc gatgggccgg 360
 atggccacaa atacacactg caccacatgc tgcgggcacc caagcatatg gctactgaac 420
 actacaggcc acagtgggtca tggggcaggg actctgggtca tagatgcagc tgagggactt 480
 gggctgggga catgtgggtga tgggtcaggg atgtatgggt agcaacatgt gttcaagagg 540
 cagtgttatg ggctagagac gtgtgggcat ccaccaggaa taagtgtttg ccggg 595

<210> 68
 <211> 580
 <212> DNA
 <213> Homo sapiens

<400> 68
 gagtcaggac ggaggacgcg gcaggtcaca gagcccacca agtccgaagc tggaagttca 60
 gattctttga tattcaaagg tggatcatct gtgctttttt ttttttatca gtctctcact 120
 ttttatccat catctaattg tgacagctta tttgccttta taccataaga tggggagtag 180
 ggttgagatg aaatccaagc atcgtttccc ttccccgatg gtcgcctccc tggggtgaga 240
 cgttcgacgt gtcagacttc accaagagca tctcccgctt cgggtgcagta atgaacttgg 300
 aaacgattta ctccggcact tggttcctgt ctccataaat gcggctgctt taaagggaa 360
 gtaaaaaggg ctgtaaattg gtattgattg ccggtgggtc tgaagaacct caactgagga 420
 ttgaccgttc cttggagtga aggctccgca ttcagacgcc tttcgctta cgtcatcata 480
 attgagaagg gaaaggagac gtgttagttt cagtctgatt atttaccatc aaggcataaa 540
 cacttctcag aggcagcgga acccattaaa ccggcccgt 580

<210> 69
 <211> 589
 <212> DNA
 <213> Homo sapiens

<220>
 <221> unsure
 <222> position is 559 nucleotides
 <223> "n" refers to an undetermined base

<400> 69
 acacgggggg caacctcttg cacctggctc cctgccctcg gtgccacgtt tccagggttc 60
 ctccacgtcg caggctgtgt cagcctcgct ccttccactg cagaattgcg gtccacagcc 120
 tggatgggcc actctccatg tatccacctg tccctccgtg gctgctgggc tgagtcgctt 180
 ctgatgctaa caagaggcgt ccggctggac taaggccccg gaagctgaga actggagggc 240

aggtgcgggc atcgggcaga gcagctccag caggcaggac ctggggcctc caccctgcac 300
 ccctgtgccc cgcgtgtggc ggaaccgccc cgaggggagg ctgtcaccac ggtgacaggc 360
 agccccacgc gagcctgaga accctcagcc cacctttttc tgtaatcaca gcaggcatct 420
 ctccggcaag tcaatccagt tccagctggg gctgcctccc ttgctcatg ggctttattt 480
 tagaactctg agcaataata aaaaagacgc taccgctac aatagatgtg gcagagaatc 540
 tggctcttca cttcatcana gatcacccctg aaatgatggg tgttggttaa 589

<210> 70

<211> 748

<212> DNA

<213> Homo sapiens

<220>

<221> unsure

<222> position is 10 nucleotides

<223> "n" refers to an undetermined base

<220>

<221> unsure

<222> position is 412 nucleotides

<223> "n" refers to an undetermined base

<400> 70

gctacatctn ctctacattc taactaacac ttgttatttt ctgtttttgt ttgtttgttt 60
 ttaatagcca ttctagtagg catgaagtgg tgtttgcctg ctttttttga tggaggtgga 120
 ggaatagggt ggaattgggc cttaaccatc aattaagctg ggggccttag acctctgtga 180
 attggctgtg acaatagcta aaggaggctg ctacctcata ctgaagagat gtttcctaag 240
 tttgtcaccc gagagggcac cgaaccaact tattgtcttg gagggaagaa gcagcaaggc 300
 agaagacttg aacttctcag agaaaaaac agtctacaga cttcatttta tgctgtcctc 360
 acacactact gaaagctcta ccctggggac ctggcttgac ttctaacctc cncctgtgtt 420
 atttaggaag agctcccagc tgctctgagt ctcagtcctc caatcagtga aatggaggca 480
 atagcacctg cctggctgca tcgccccaca gtgctgcaat gagcatccaa cgagagaaag 540
 cttgtcacct gtgttgcaaa ctaagttaca caaatgcagg cagtagcagc tagaagaaaa 600
 tggttgggaa tctgaaaaga attaaagccc cccatgaatt tcttctcacg ctcctccaa 660
 aagccaggga ctgcttcacc ccgcctccag gactgctcgc tccagcattt ccggcagctg 720
 ctgacagaat gtatgttgcg gctgtccc 748

<210> 71
 <211> 599
 <212> DNA
 <213> Homo sapiens

 <220>
 <221> unsure
 <222> position is 491 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 522 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 538 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 584 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 596 nucleotides
 <223> "n" refers to an undetermined base

<400> 71
 gatgactggt gcccgagctg aggccacgac ccaacccga ggaagggaga acagcttccc 60
 atgaagggca tggctgctgc ccataatcc cagggcagga aataaaggga tcttggacta 120
 ggcaatcaaa ggacttcctc tccctctaag gccaaaggagg aaatgtggct gggactccaa 180
 gctctgtgga tgcttgagg tgccagcagc tggggatcag ctggccccac ctgcagagcc 240
 agccagtggg cccctgcat ctccaagggt gggcttatgg gctccaagaa caggtgtttc 300
 tcagggtaac ctgagccct acaacttcaa ccaagagagt gaaggggagc agccctggag 360
 gccaatgagg agggggatta gtggtcactg atgacaaaga catccctgtc ccagagacca 420
 gcccttgtg agcagaagaa tggtgccgg gcaaaaggac ctgctatgcc ctccccatac 480
 acatatcatg ncacctgggg accctctgaa taacaggggg cngctttaga gtggcttnat 540
 taccaacaag agggccagaa gggctagagc acacgatttc atgntcggcc gcatgncaa 599

<210> 72
 <211> 614
 <212> DNA
 <213> Homo sapiens

<400> 72
 gtgcgctatc acgactgttg cccgagctga ggccagaccc aaccccagg aagggagaac 60
 agcttcccat gaagggcatg gctgctgcca ccataatccc agggcaggaa ataaagggat 120
 cttggactag gcaatcaaag gacttcctct ccctctaagg ccaaggagga aatgtggctg 180
 ggactccaag ctctgtggat gcctggaggt gccagcagct ggggatcagc tggccccacc 240
 tgcagagccc agccagtggc tccccctgca tctccaaggt tgggtctatg ggctccaaga 300
 acaggtgttt ctcagggtaa cctcagcccc tacaacttca accaagagag tgaaggggag 360
 cagccctgga ggccaatgag gagggggatt agtggctact gatgacaaag acatccctgt 420
 ccccagagcc agccccttgt gagcagaaga atggctgccg gggcaaaagg acctgctatg 480
 ccctcccat acacatatca tggcagctgg ggagccctct gaataacagg gggcgcttta 540
 gagtggcttc attaccaaca agaggcccag aaggggctag agccacacga tttcatggtc 600
 ggccgcatgc gcaa 614

<210> 73
 <211> 552
 <212> DNA
 <213> Homo sapiens

<400> 73
 aagcgccac agatggccaa gcatgtggag gagagcaca tttttatatt aaatatccaa 60
 atacgaacac attcccgcat ggcaccaaca gccgcctgaa cagccccgat gccggcttgt 120
 gctttttccg ttttgtctag aaatttgggt tgcactaaat tctcagctga atgaagatga 180
 gaaggggctg gcagaggggg tggctccagc tctctgagaa cctggctcct tcccgggtgg 240
 cagggagaga tggccccctg ggagacgggg aggggtgact gcctcatgcc caaaccacca 300
 gcttctagtt gagaaatcag aattttctct gcagaataag gaaaaagcat tgtcaccatg 360
 attcacgtgg agctggccac actcaggaaa ttcaatgggg tcccacaggg gctccgaggg 420
 ggaaggagag ggctggggac atgcccctcc agccatcatg gaacaggatg ggcagggccg 480
 gccctcactg ctctctaaca gtgaaaagcc acatctccac tttggaaaac acaggcatgt 540
 gagagcctgg gg 552

<210> 74
 <211> 450
 <212> DNA
 <213> Homo sapiens

<220>

<221> unsure
 <222> position is 378 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 403 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 409 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 440 nucleotides
 <223> "n" refers to an undetermined base

<400> 74
 tggaggcttc gagggaagtg aggttccttc ggacacccta gtgggaaggc tccacgcggt 60
 aatggaacca cgctgtgaaa cctttgcctt tgggtgtcat ggtggaagca aatcttagaa 120
 gacatttaat ttaaaaaatt cagttttaaa aaatgttgac ttaaaaagca gttttgaaaa 180
 acaacctgga attagcctga gatcgatgcc aactcttagc agtctgtata ctaaacacag 240
 ttaaacaact gtagctgctg gcaagctgga acctttttgt aaagaagcac ataaaaagga 300
 cagaactggt ggaagggtgca ctggtctttc cacatcgcca ccaggcggtt tgaagcgtgc 360
 tgctgacacg ctactcanat gcttctggaa gccaaacaat aanaaaaaanc cccattgttt 420
 cccttgctgg gttttaccn ccatggtgga 450

<210> 75
 <211> 432
 <212> DNA
 <213> Homo sapiens

<220>
 <221> unsure
 <222> position is 417 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 421 nucleotides
 <223> "n" refers to an undetermined base

<400> 75
 ggacaatgag gagggggtgc acgtggaatc cccacggata ggccggacgc cgggcaggag 60
 cctttgcagg ggtgcacagc ctctctgga agccctggtc gctgcctggt gctgctgca 120

```

ccctgcgggc tccgcagcgg tggagccagg cctgaactgc ctgctcttgg ccccgccctgc 180
ggccctctgc cctttgtctt gcccggtgggg cccggggcct caagctggcc cggggttcct 240
gaagttagct gacgatgggc tggcctctgg ggctgggtcg tgggccttgt gcaactggccg 300
ccacgtcacc agcgccaggc ctaccgcggg tgctgctgga gacgcgggat gcccgggctc 360
gggctgtgct ggatccctg gcgctgcgaa ccccgtagcc ctttccaatc gcgggcnccg 420
nttaaagccc ga 432

```

<210> 76

<211> 501

<212> DNA

<213> Homo sapiens

<220>

<221> unsure

<222> position is 18 nucleotides

<223> "n" refers to an undetermined base

<400> 76

```

gacgagacct agccggcncc atgcgcgcct tgagcctggc gaacagttcg gctggcgcgga 60
cgcgcctgat gctcttcgtc cagatcatcc tgatcgacta gaccggcttc catccgagta 120
cgtgctcgct cgatgcgatg ttctgcttgg tggcgaatgg gcaggtagcc ggatcaagcg 180
tatcgagccg cccgattgca tcagccatga tggatacttt ctgggcagga gcaaggtggg 240
atgacaggag atcctgcccc ggcacttcgc ccaatagcag ccagtccctt cccgcttcag 300
tgacaacgtc gagcacagct gcccaaggaa cgcccgtcgt ggccagccac gatagcccg 360
ctgcctcgtc ctgcagttca ttcagggcac cggacaggtc ggtcttgaca aaaagaaccg 420
ggcgccccctg ccgttgacag ccggaacacg gcggcatcag agcagccgat tgtctcgttg 480
tgcccagtca tagccgaatt c 501

```

<210> 77

<211> 826

<212> DNA

<213> Homo sapiens

<400> 77

```

gcgcccctgt gggatgacgc accatcctgt ttgtttgcac caagtcattt atctcgtgca 60
ccccaggggg ccgtgggtccc tgccggggcca tcatgtctgc ttcccttatt tgggttttct 120
gccccctcac ttcattttctc acttcgcttt tcctccttat ccttttgcag tcttgctttt 180
gggggcattg ctacgccagt aatttgaggg acacctcgtg gagccctagt gtggagccgt 240

```

cagagcctgg gtaggattct ccgtggtgag gtgctcaggg agacacagga gcattccggc 300
 gcctgttcct tgtgcacatc cgcaagtgtc tgcagtgaga ggcatgggtc ccatcttgaa 360
 tgccaacaat gtggcaccca caccctactt gatggggccg agccacagct ggccagggtg 420
 accaccatgg acgtgccaga ggcacccgaa acccagctct tgcccagctg ttccactgcc 480
 aactccagcg ttagcaaagc agctctccct tgctttgtct tctacagcag agaacagatt 540
 aaaagagaag ctgcaggcag agaaatgcct cttggagcca gatgccccaa aggatctctt 600
 tgaacaaagg gttgctcagg tcagcgttag ttcttgcat caagcaaaa aatcagagat 660
 gctaacagtt ctacagattca ctccaagtga agactcaaag ctggatttat aaatccccac 720
 agagccgctg tgcagaggta gagggccggt ttcaggatga ggaagccctc ttggaagcac 780
 cgtcctccgg ctaacaagcc tccaacctac tgcggcagg gagaac 826

<210> 78
 <211> 433
 <212> DNA
 <213> Homo sapiens

<220>
 <221> unsure
 <222> position is 16 nucleotides
 <223> "n" refers to an undetermined base

<400> 78
 tgcgcagctc cgcgangtgc ccggcggggc cgaccctcag actcgcttgt ccctggagac 60
 caaccctagc gaccaggctc tgccggatcc cgtcggggtt caactcctat tccgaaggctc 120
 ctttctcccc taatcacaac acccactcgc ctctttttcc tctcttctc cagcttccac 180
 cgccgaccgg gcagccccag ttacccgata acggctccca aggccccgtg ttacattct 240
 ttccactgg aagcagaaat tatcagccc aaattcctac ctgccttccc tggattcctg 300
 gtttcctaag aaacggggtt ggcccacccc tgggcgttcg aacagtccac agaagcgggc 360
 aaaggaaaga cgactcagtc tttccctccc gccaatctct tctccgggac cacagatccc 420
 agaagtcacc gcg 433

<210> 79
 <211> 424
 <212> DNA
 <213> Homo sapiens

<400> 79
 ggccggggccg accctcagac tcgcttgtcc ctggagacca accctagcga ccaggctctg 60

ccggatcccg tcgggtttca actcctattc cgaaggtcct ttctccccta atcacaacac 120
 ccactcgctt ctttttcctc ctcttctca gcttccaccg ccgaccgggc agccccagtt 180
 acccgataac ggctcccaag gccccgtgtt tacattcttt cccactggaa gcagaaatta 240
 tcacgcccaa attcctacct gccttccctg gattcctggt ttcctaagaa acgggttttg 300
 cccacccctg ggcgttcgaa cagtccacag aagcgggcaa aggaaagacg actcagtctt 360
 tcccctccgc caatctcttc tccgggacca caaatcccag aagtcaccgc ggccgctaag 420
 ccga 424

<210> 80
 <211> 285
 <212> DNA
 <213> Homo sapiens

<220>
 <221> unsure
 <222> position is 14 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 27 nucleotides
 <223> "n" refers to an undetermined base

<400> 80
 caaccggggg gcanaggcga tcaaaantgg ggtgcgctgt ggtgggacgac acgtgtggcg 60
 cgggtctcat tatccgccct ttccacttcc tggactggaa atggcagacc atatgatggc 120
 aatgaaccac gggcgcttcc ccgacggcac caatgggctg caccatcacc ctgcccaccg 180
 catgggcatg gggcagttcc cgagccccc aaccaccag cagcagcagc cccagcacgc 240
 cttaacgcc ctaatgggcg agcacatata ctacggcgcg ggcaa 285

<210> 81
 <211> 401
 <212> DNA
 <213> Homo sapiens

<400> 81
 cagatatgta tcctcctctt tccaaccctg cgtccctttg aggcctggtc ggcgttccca 60
 acctgcccct accccaccaa cccctgtccc tttggccatt agtcccggat tatctagcga 120
 tgcccctgt accgtctggc tttgctgttt actccgcgct cggccagttg aggccttttg 180
 tttttattcc tgattttctc ataggggtaa agtgccctcg ggaggatagg acaagtccca 240

tctgttcat acgaattaca gtcggactt cgggcccttt tacactgcct tttgtatctg 300
 ttaacttgcg ctaaaaacga ttcggttctt ttttttgagg aaggggggttg gggggcgag 360
 actctgtcgc ccagtcctga gggccgcggc gcgcaagccg a 401

<210> 82
 <211> 268
 <212> DNA
 <213> Homo sapiens

<400> 82
 atagcgcgca caactgtgtc tcttaccag gcacatgcac tatccctgat cccggtgcat 60
 gatgggaatg tagtcctgca gccctgtgac caaagggctg ggagtgttta tgagacagca 120
 tctctcagca agcaaagcaa ggctgcaca gcccgcctt ttcctccagt gaggcgcact 180
 gttcattaag gagtgttcat gagattacat tttccatcaa gccagccag tcacgcacag 240
 ctctacctct tctctgccg ccccgcaa 268

<210> 83
 <211> 989
 <212> DNA
 <213> Homo sapiens

<220>
 <221> unsure
 <222> position is 878 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 884 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 918 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 929 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 973 nucleotides
 <223> "n" refers to an undetermined base

<400> 83

```

gggtaatggg ggtgaacaga gagggatgcc gaggccagct tgtagtgtgg ctgttgggtct      60
tgtccatcct atggcacaac cctgtcacca cccagatttt gttaggagtc ctcccccaac      120
ttgagagtgg aagctccttt ggcacaaaaa ggggttctgc atcatcccc agccccagc      180
cctgagcctg ggtctggctc tgaactagac ctccatgaat gaatgcacag catcagtggg      240
gatccacat catggggaaa tagtagatac aggaatgatt ttccaaccag attacagact      300
atttcaagcc cagccagagc ctaccaggcc aacattcccc aggettgtgc ctctccgagc      360
ctcagattgc tcctcttca aacgagggac agctctgctg gcattacctg aactctaggg      420
tcctttataa gtcagactc cagcttagag cacacattga gaggctgctg caccacagag      480
ccacatacgt gcaacagagg gtggtccaga ccccttattg gtcccatgg ggtttgagag      540
agaagcctcc agaccagctc aacttctccc tcctctcact taggcctttg caccagctc      600
ttaggagggt gtcagggtac agtgcccat ttcttttctc ttcccagaa atcatgcggg      660
ggataacctg tcagacagga ccttcatgaa agccaggctg tgagggtgtg tggggaatgc      720
ataattgata ggccatcgtt cggaggccct cctggaggac caaaatgtaa tcagcagtgg      780
cgagcttggt cagcagagga attcctttta catcctggtg aggccaaaga cctggcaagc      840
aagtcctct ggtcattaaa gaagcatcct gacttgangc aggnacactt aggtcactgc      900
agccacaaaa atctttgntg ctggattcna aagtaggcat tggggctggg atctgggctc      960
tggcatcctt gancgtgtcg ggggccaaa      989

```

<210> 84
 <211> 250
 <212> DNA
 <213> Homo sapiens

<220>
 <221> unsure
 <222> position is 37 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 40 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 49 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 75 nucleotides
 <223> "n" refers to an undetermined base

<400> 84
 cgggctcgaa acttcgaaga ccgcggaacc cgaagcngcn cttggctcna atcgcttcgg 60

ctcgaggcgc ccgtncgggt cacgtgaggt gggggcgggc cgaagagggg ggctcccctc 120
 ctctgccgc agggttggcc gcaagtgcgc ttcaagaggc gcttgatgac ggttaatgtt 180
 gcagcccgga agatgacttt tttctcctcc ttgggttgcg gcaggccgtt agtgggaggt 240
 cgcgccccga 250

<210> 85
 <211> 402
 <212> DNA
 <213> Homo sapiens

<220>
 <221> unsure
 <222> position is 224 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 265 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 382 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 390 nucleotides
 <223> "n" refers to an undetermined base

<400> 85
 ttctcccttg tcatcccctt accagagcca cagaaattat ccctgtgggc tcccttgtcc 60
 tcaactcggcc ttttctggag ttaagagatc caagccaact actgggtctg ttccctgcta 120
 aaatcttagg ccggcgtccc atccacccat ccccatgcct aggactttta agctggcaac 180
 ggtacctggg tttagttttc ccttcgtata tcaactatctt cgtnngcttac cttcttgtgc 240
 ctaaagttcc accgatgtgc aaggngatta accactaaag tgcaacctgac actactcttg 300
 acaaattgca gttgggaggt gagttgatga ctggccggta aatcaaaaagt gcttatttag 360
 ggagtgaggg ggcccgcggc anaagccgan ttccagcaca ct 402

<210> 86
 <211> 595
 <212> DNA
 <213> Homo sapiens

<220>
 <221> unsure

<222> position is 157 nucleotides
<223> "n" refers to an undetermined base

<220>
<221> unsure
<222> position is 377 nucleotides
<223> "n" refers to an undetermined base

<220>
<221> unsure
<222> position is 410 nucleotides
<223> "n" refers to an undetermined base

<220>
<221> unsure
<222> position is 441 nucleotides
<223> "n" refers to an undetermined base

<220>
<221> unsure
<222> position is 444 nucleotides
<223> "n" refers to an undetermined base

<220>
<221> unsure
<222> position is 456 nucleotides
<223> "n" refers to an undetermined base

<220>
<221> unsure
<222> position is 461 nucleotides
<223> "n" refers to an undetermined base

<220>
<221> unsure
<222> position is 473 nucleotides
<223> "n" refers to an undetermined base

<220>
<221> unsure
<222> position is 490 nucleotides
<223> "n" refers to an undetermined base

<220>
<221> unsure
<222> position is 525 nucleotides
<223> "n" refers to an undetermined base

<220>
<221> unsure
<222> position is 532 nucleotides
<223> "n" refers to an undetermined base

<220>
<221> unsure
<222> position is 534 nucleotides
<223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 541 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 572 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 575 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 583 nucleotides
 <223> "n" refers to an undetermined base

<400> 86
 gatcccagaa ggttctggag ccgagtatca gagtttgagc agcgagtcca gcctagcaga 60
 agcgggtgtt gaccggagac ttttcaatgg tgcaaaatga cacactgctt ttgacttggg 120
 gatctgtccc ttgtggcacc agaagctaca acaggtncac ctggattcca gctctagctg 180
 gactcggtaa ttgctaagtg ccagctctga agtctgtgat tccgtggaaa tccctttcaa 240
 gccgaattc tgttttttat gggcctcttg tccaaacagt ttgacttgtg aactctgttt 300
 ctgtcaagtt gacacttggg ctgggcaccc attcatgagc cagatgaaag cggctaaatg 360
 cccgaaaaaa taaaggnttt tacctttttt ttgaaccatt ggtgagcatn taaaaaaatt 420
 agggaaggta aaaccaacc nggncaaacc caactnaaca nttttttttt ccnaaacaag 480
 ggggggctan tttttcactt ggaaaaacaa acaattttta ttgantcttg ananggtgga 540
 naaccaaatt tttttgttgg gttgggttcc gnagnccgaa ttntgcaaatt ttctt 595

<210> 87
 <211> 304
 <212> DNA
 <213> Homo sapiens

<220>
 <221> unsure
 <222> position is 279 nucleotides
 <223> "n" refers to an undetermined base

<400> 87
 cgtggcccga tgcattcagg gagccctctg tgttggccgc atagcaggtg tagttgccgg 60
 catcctggat gaagacgggc gcgatctgta gacccccga ttcaagaagc atgaacctag 120

gaatccggac agagccactg gccagaatgt ggTTTTctaa agaacagtgg agaaaagagg 180
 catgttacag tcgtaacgct tgaaggaaat gaagatagt gttagagcca taagcaagta 240
 atatggttcg gctccgtgtc cccaccaag tctcgtctng aattgcaatc cccacgtcgg 300
 cgca 304

<210> 88
 <211> 296
 <212> DNA
 <213> Homo sapiens

<220>
 <221> unsure
 <222> position is 9 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 31 nucleotides
 <223> "n" refers to an undetermined base

<400> 88
 ggctttcgnt aggagttaat ggggcattgg ngggtgggat ggcagggctg ccagcatctg 60
 acccaggagg ctgggaggag gctgctgtgt gaatacacgc tcggcctctc acagtggctg 120
 ccgccgcatt agccccttgt gtttcaggga acagagcatc cgtgatggat gagacttta 180
 ttaaagtaat gagacattta taatcgcggt tatctccaaa attaggcctt ttagcaatta 240
 ttctgtggga atattctctc ggtagatagc tcccttttta gaacaacgtc ggcgca 296

<210> 89
 <211> 220
 <212> DNA
 <213> Homo sapiens

<220>
 <221> unsure
 <222> position is 10 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 24 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 29 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 30 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 31 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 38 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 45 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 87 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 99 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 134 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 158 nucleotides
 <223> "n" refers to an undetermined base

<400> 89
 attggcccg n caggcgggaa acangctg n nttctctnac cgttntccag cactgcccag 60
 accaggaggc gcagggagag gaggggncag cggttccgng accgctcctc ccgctgtccc 120
 tgctctccag cctntgcctc tgcaggagcc cgcgggan tt gccccaggcc cctgtcccca 180
 cctgtggctc ccgtcctggg cgctcccggg gccgcggcaa 220

<210> 90
 <211> 273
 <212> DNA
 <213> Homo sapiens

<220>
 <221> unsure

<222> position is 2 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 7 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 10 nucleotides
 <223> "n" refers to an undetermined base

<400> 90
 gnagggnggn ggtcgcgac gccggtgggc agttcttggt cggtgatgtg ggttaaaaag 60
 gactgcagcg aggagccggg gcggcgctcg gagtaatcac cggcggcatc aaaaagcgcc 120
 atcatggcat cgaggtcgcg gtctgcttgg gagccggtgg cgccgccgcy caaggcagat 180
 gcctgcaggc gcatatccag ctcggtagcg ctccatacct cccacaggat ttcttcaca 240
 gaggttggg cttgtatagc ctgccgccc gca 273

<210> 91
 <211> 361
 <212> DNA
 <213> Homo sapiens

<220>
 <221> unsure
 <222> position is 10 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 12 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 212 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 218 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 356 nucleotides
 <223> "n" refers to an undetermined base

<400> 91
 acggcttctn tnctaagtga cacggtgtgt gaaattcggg tggggaggta gttctgtaaa 60

ctgcgtctcc ccgccagcta aggaagttga gtgaaggag cggtgccgtc tgggaatcgt 120
 agtcctcaca aaggcgtgag taggcggcaa ataaggattt gggtttagcc ttggggattc 180
 actcctgtca aagctgttag agaagctccc anaactcnta aagtaacaga aactacttgc 240
 ggcaacattt gtaacttcca cctggctcat tatcttcac tgttaccttg tgttctagat 300
 aagttataat ttattctaca tatcgttcag aagtcttggt cctgttccat attgtnagca 360
 t 361

<210> 92
 <211> 462
 <212> DNA
 <213> Homo sapiens

<400> 92
 gctgcccaca ctggatggga aggaccggcg cctgcagcat ctgccctcca agccttcgta 60
 gctccctcct tcctgcagga taaactctaa actccttagc acaacgtggg agccttctca 120
 gagactgggt ccaaccctac tccagccgca gcctccctc ctggcccccac tgccacaccc 180
 ccgggcctcc ggccacactg agcctctccc ggtttcccag gatacaacac tcgcccattc 240
 atagtgtggt gccttttgca cgtgctgttc ctctgcttgg ggatgctgtt ggtctttctc 300
 agccagggtga agaggacgct gaatgtcacc tgcttgagta tcaggaccgg ggactgggag 360
 ctggacctag actcttggcc ctggagagaa gccctgcatg gggccgcagc ctgcccccg 420
 ccctgctcac agaaaagctc agccttgagc ccgcgtggga ga 462

<210> 93
 <211> 591
 <212> DNA
 <213> Homo sapiens

<400> 93
 caaagtcacc tccacgggtg ggctcagcag ctccgcacac ttggtcatgg tgcggggaa 60
 ggcccccctc agctgtaggt gggtagtggc agaacaggag ggtgagggga gagtccgaac 120
 tgtccccact tggccgttcc ctccccactg gggggccctg agccagtggc ctctctctc 180
 ggggcctccc cggaaggagc caaggtctgt ctgcgaggca ccggtccccg gccacggcca 240
 tcagccccca gaggtggatc agggcatcac cccactcca cagctgaggc caggggggtca 300
 gggaggcaac cagggcagac ctggaacctg gctctgagac aggacggccg agggcccctc 360
 cactctccct ccctcggggg gggcactgac ctggacgcca aagatgtcct cacactgggtg 420

gcgtttgagt agggcccact cggacatctg gccctgcagc aggttggtgc agacggccat 480
 ctctccacat gtcacatccg ccccgaagcg cttgcagatc cgtcggaagg gcaggttccc 540
 aactgcggg gggagcagga cagacacaca tgctcttgca cgcgcacctc a 591

<210> 94
 <211> 279
 <212> DNA
 <213> Homo sapiens

<220>
 <221> unsure
 <222> position is 3 nucleotides
 <223> "n" refers to an undetermined base

<400> 94
 ttntgagttt tggcctgccc acagtctagc cctggacaga gaatccgagg ctccagccatg 60
 ctgcagcacc caggacactg catcccagca cctgcccga aatcagccca gggacccaaa 120
 ggaaagcagg ctccaagctc cccggaagcc aaggaaaata ggaaaacata tcctgccccg 180
 gggacacctt ctggaactat gaccacatgc acttgacctt ccggaacaat caccgcatgc 240
 acctgacctc ccggaactgt caccaccgag cgcacctca 279

<210> 95
 <211> 351
 <212> DNA
 <213> Homo sapiens

<400> 95
 cctttattat tggttaaact caccagaaa acccttaact cttagacagc ggctctcatt 60
 aagcaaaagg ggaggcacat gaagctccag gcagggcccg gagggaaccg tgaagccaaa 120
 ggctctggga gccccaggc acctgcgttt gcattttcat cctggaggag accaggcctc 180
 tggggctgct ccccggggtg cagagaggag gggcttttct tgggtgtgaa catactcatt 240
 gattcagtca cctgaccttt gactccatgt atttgttga gtctggatgt gtggtgtgct 300
 ctgcccagca gctgggatcc acatgagcac agacatggtc cccccgcggc a 351

<210> 96
 <211> 171
 <212> DNA
 <213> Homo sapiens

<400> 96
 ttgagtgtcg cgtgaatacc taggggacac tcaggggaat gatggctccc ccgagaggta 60

aaggggtggaa agaagggggcc tcagcaggtt aggtcttctt gggtccttct gtagggcgtc 120
tgggagatag atccgtgggg ctcctagggt cgcccctacc cggcgcgggc a 171

<210> 97
<211> 743
<212> DNA
<213> Homo sapiens

<220>
<221> unsure
<222> position is 155 nucleotides
<223> "n" refers to an undetermined base

<220>
<221> unsure
<222> position is 181 nucleotides
<223> "n" refers to an undetermined base

<220>
<221> unsure
<222> position is 202 nucleotides
<223> "n" refers to an undetermined base

<220>
<221> unsure
<222> position is 228 nucleotides
<223> "n" refers to an undetermined base

<220>
<221> unsure
<222> position is 259 nucleotides
<223> "n" refers to an undetermined base

<220>
<221> unsure
<222> position is 262 nucleotides
<223> "n" refers to an undetermined base

<220>
<221> unsure
<222> position is 293 nucleotides
<223> "n" refers to an undetermined base

<220>
<221> unsure
<222> position is 366 nucleotides
<223> "n" refers to an undetermined base

<220>
<221> unsure
<222> position is 386 nucleotides
<223> "n" refers to an undetermined base

<220>
<221> unsure

<222> position is 388 nucleotides
<223> "n" refers to an undetermined base

<220>
<221> unsure
<222> position is 447 nucleotides
<223> "n" refers to an undetermined base

<220>
<221> unsure
<222> position is 470 nucleotides
<223> "n" refers to an undetermined base

<220>
<221> unsure
<222> position is 484 nucleotides
<223> "n" refers to an undetermined base

<220>
<221> unsure
<222> position is 502 nucleotides
<223> "n" refers to an undetermined base

<220>
<221> unsure
<222> position is 512 nucleotides
<223> "n" refers to an undetermined base

<220>
<221> unsure
<222> position is 516 nucleotides
<223> "n" refers to an undetermined base

<220>
<221> unsure
<222> position is 590 nucleotides
<223> "n" refers to an undetermined base

<220>
<221> unsure
<222> position is 664 nucleotides
<223> "n" refers to an undetermined base

<220>
<221> unsure
<222> position is 667 nucleotides
<223> "n" refers to an undetermined base

<220>
<221> unsure
<222> position is 673 nucleotides
<223> "n" refers to an undetermined base

<220>
<221> unsure
<222> position is 695 nucleotides
<223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 717 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 742 nucleotides
 <223> "n" refers to an undetermined base

<400> 97
 cctccctggc ccttggtccc aaggagcttc ccttggtccca gcctcttcgc cagtgaacttc 60
 tcaactggacc attcctttac aaggagcctg ttttttgtgt ttttttttta cacctttttt 120
 cttctatttc acagaaggaa caccggacgt ccctntgtga tggcagcagc catgctgcct 180
 ntgtttccgc tcaggggttc tntgccacct ccaattccac ccagtctntt ggctctggct 240
 gggcttcggc tcccgccnt gngccaaaaa ttgcaatgcc cgcggtcagg gcnctttgcg 300
 gagtctcacc gcctgaggag gcttgattcc ctctcacag gcagcagcgt ttgatggccg 360
 gtgacncccc cttttccaag cacatntntc atggccctg aatgccactt acagggcgtc 420
 cctccctgtg ctaagtgtg cctgganctt tgggtgtggc agcagcaaan acctctaccc 480
 ttngngatgt tcgtttcggg gnggaaagac anatancaaa gttggtcgta aactgtaaag 540
 tgtgctggga ggaaactgag gcagggaggg cctggtgcca ctggggagcn ctgccccgac 600
 cccatgtgct tcccaggctc ccttgagacc acgtggatgg cgacttcctg accttgagg 660
 ccgnggncct cantcctcat gctcgatggc gtcanccccc tcttggggaa atccaancat 720
 tcttgacctg aaaatgcacc cnc 743

<210> 98
 <211> 589
 <212> DNA
 <213> Homo sapiens

<400> 98
 ttgccgcgct gataaaggaa gcgtctagaa ggtctcccca gccttcatca tctgagactt 60
 ggctttcagc cccaaagcac taggcctgc tgttaacctt ccaccattaa cttttggtgc 120
 tcttcaatta gcagcagcca ggggtccttg gcaggtatga gaatttgga ggacagcccc 180
 agggcatggc ccccggtgc agcaaaagtt ctaagtgttc ttctgttgga aggaagccca 240
 ggagatattg atcagctgca ggtgggggag gccccagatc ccaccttgc ctgcctccag 300
 gagaagggtc tccatgggcc aaaatggagg cagagtccca ctttgcctgg gcagctccct 360

gagcatggct ccctgtggac ggagctgagt gacgtcatga ctctaggcct caacaaaaga 420
 gctttggaaa atccccgatga ttcgaattgt attaaatcaa caaacatcgg gttgcacagt 480
 tactagaaaa cggagatctg cgtcatcact tactagacac gtgaccttga acggcggcctt 540
 ccccggtgtga aacagcaaag ttctgtaacc cccatgaacg cgctctca 589

<210> 99
 <211> 538
 <212> DNA
 <213> Homo sapiens

<400> 99
 tgccgcgtct gaccctactc tcacaaagac tttccaacta gcataattga gttaaatggt 60
 ccccccaact cccttaattc aagctaaact tgcagtttaa caactatagg agtgatatct 120
 acacattaat gccacacttt aacatgccta acactacaca tgaacacgct tccgggtgct 180
 gttacatccc gctctctccc aagcacgaga cacaggcagg atgctgacgt cctgcttctc 240
 tgctgcgggc gggaagtcaa gactccggat ttgctgcagg agttgccgtg gggatcctga 300
 cttcacgcag gagatggteg gcctctggaa gtgcctggcc cgtttatcct tgaaatctac 360
 ctgtgcaggt ggtccttgcc tcagcccctc aggacaacac aggtctttcc taagttacag 420
 ggagaccatc agattgtcgt gtccgagccc cctgaagtgg aaccacagt ctccattcag 480
 tctgccctca gtttccctcc cctctgcagg gccattgctg ctgtggacgc gcctctca 538

<210> 100
 <211> 486
 <212> DNA
 <213> Homo sapiens

<400> 100
 agaggtagaa aaaggagtta gaagcaaaga ggaaaaata aataaacagg caacaaaaac 60
 ccaaccagc cagcctgagc catttgcat agtggtcatt taggaaatta gcagacggga 120
 aacgctgggg agtggagtgg gccccggcct tggggactgc agagcccgct cagccctggg 180
 tggctgggcc cacatgggct gtgccccagg agcacaggag gacccagagg gtggccgaga 240
 gagcctcgcc gggctccggt atgggtcctg gccctcaca ggtgcgagcc tggcccagtg 300
 actgtggacg ctgtgggaga gcaggcctcc gatacgagg gctgggactg ctgacctgga 360
 aggtggtgcc gggcgtgtct ggtgaaggcg ccgttggcag ctagagagag acggcggatg 420
 gggtagcgcc attaccacg gtcccagttt tgaggcttga cggtgacgga aaaggacgtc 480
 ggcgca 486

<210> 101
 <211> 450
 <212> DNA
 <213> Homo sapiens

<400> 101
 aattgaacca ggggtgcacgg ccagcgccag acacagtgag cttcatggca actccagttt 60
 accggtgaga accatggggc cactcagaga ggcaaagagc ctcacccgag tgagtcctct 120
 ggcttctccc cacctggggc gggccccagg ccgcgctgtg gttccctttc cagccgtcat 180
 ccctgggtga tgggaggtgg gcattctgtt caaccttgtg ggtcagggag ccagggccag 240
 tgtgcagatg agaagaggct gcggttactg gcgatgagag ggactgtccc cttcgtgggc 300
 actttctctt ttgaggccag tgaaatgtgt tccctggggg tgtattcctg agaaggcctc 360
 atttaaaggg agccgccaaa ccaagtgggc ttagcaaaag cagtttgtca cctggcagca 420
 cgtgtgagcc tcgcccggac gcgcctctca 450

<210> 102
 <211> 292
 <212> DNA
 <213> Homo sapiens

<400> 102
 agcgcggcct ggcagattgc ccattaatga aactcagtgg gcagaggctg ctgagggaca 60
 cggattccca ctccccgggg gagggggtgg aaatggcttc ctccctctgc ttccctacca 120
 ccagtaatgg ggagctcacc atgcttagaa gactcttctt tgcattggagt tcgggcctcc 180
 tccctgcacc taccacccta gtggcccaa gtcttaaggc tgaagggttaa tcctgtgtcc 240
 ttcagaagca aaggctgcaa ccgataccaa acagaggtgg ccagcgcggg ca 292

<210> 103
 <211> 395
 <212> DNA
 <213> Homo sapiens

<220>
 <221> unsure
 <222> position is 340 nucleotides
 <223> "n" refers to an undetermined base

<220>
 <221> unsure
 <222> position is 367 nucleotides
 <223> "n" refers to an undetermined base

<400> 103
agagcttata ccgcgagcac aagggagccg gggcctgggc cgccgtggga aggggctcct 60
gccttccggg gacgcggtca gggaaagtcca gccgggggtgc tctctgcaact gcgggtgccg 120
ggctcggcag aggccaaaccc ggcaaaacga gcaggatctc ccggccccac cctagtgggc 180
tccgcctgcc ccaacaacca tcctgccatc ctccctgcga gacaggtgac tttcctctct 240
gatgcgggtgc atctgtcatc tgtctaacgg gccattccc cagtgaacaa cccccaacca 300
aagacacgaa ggggaaggcg caagcttcta ccaagctcan ttgccccatc tggtgccac 360
ctgcctngta ttggtgact tggaggatag gaagg 395

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
17 October 2002 (17.10.2002)

PCT

(10) International Publication Number
WO 02/081749 A3

- (51) International Patent Classification⁷: C12Q 1/68, C12N 15/11
- (21) International Application Number: PCT/US01/51652
- (22) International Filing Date: 26 October 2001 (26.10.2001)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
09/699,243 27 October 2000 (27.10.2000) US
- (71) Applicant (for all designated States except US): **UNIVERSITY OF SOUTHERN CALIFORNIA** [US/US]; 3716 Hope Street #313, Los Angeles, CA 90007-4344 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): **MARKL, Isabel** [US/US]; 1005 Rashford Drive, Placentia, CA 92870 (US). **JONES, Peter, A.** [US/US]; 4645 Lasheart Drive, La Canada, CA 91011 (US). **TOMIGAHARA, Yoshitaka** [JP/JP]; 2-10-2-246, Sonehigashi-machi, Toyonaka, Osaka 561-0802 (JP). **LIANG, Gangning** [CN/US]; 3436 Ashbourne Place, Rowland Heights, CA 91748 (US). **FU, Hualin** [CN/US]; 500 Norht Atlantic Boulevard, Apt. 310, Alhanbra, CA 91801 (US). **CHEN, Jonathan** [—/US]; 1008 South Marguerita Avenue, Apt. 1, Alhambra, CA 91803 (US).
- (81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- Published:
— with international search report
- (88) Date of publication of the international search report:
4 September 2003
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: DETECTION OF ABERRANT DNA METHYLATION AS MARKER FOR HUMAN CANCER

(57) Abstract: There is disclosed (103) novel methylation-altered DNA sequences ("marker sequences") that have distinct methylation patterns in cancer, compared to normal tissue. In many instances, these marker sequences represent novel sequences not found in the GenBank data base, and none of these marker sequences have previously been characterized with respect to their methylation pattern in human cancers including, but not limited to those of bladder and prostate. These (103) sequences have utility as diagnosis, prognostic and therapeutic markers in the treatment of human cancer, and as reagents in kits for detecting methylated CpG-containing nucleic acids.

WO 02/081749 A3

INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 01/51652

A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 C12Q1/68 C12N15/11

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C12Q C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

BIOSIS, EPO-Internal, WPI Data, PAJ, MEDLINE, EMBL

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE EMBL [Online] retrieved from EBI Database accession no. AL355593 XP002227132 97.576% identity (98.773% ungapped) in 495 nt overlap (498-4:167170-167664) with SEQ ID NO: 1 abstract	1,2,4-12
X	WO 00 01816 A (IMP CANCER RES TECH ;KNOWLES MARGARET (GB); HABUCHI TOMONORI (JP)) 13 January 2000 (2000-01-13) see e.g. page 2, lines 34-37; page 6, lines 14-29; page 30, line 32 to page 32, line 30; claims. --- -/--	1-12

☒ Further documents are listed in the continuation of box C.☒ Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"Z" document member of the same patent family

Date of the actual completion of the international search

15 January 2003

Date of mailing of the international search report

02.05.03

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Rojo Romeo, E

INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 01/51652

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 5 552 277 A (LEE WEN-HSIANG ET AL) 3 September 1996 (1996-09-03) see e.g. column 2, lines 16-29; Examples and claims. ---	1-12
X	SALEM CAROL ET AL: "Progressive increases in de Novo methylation of CpG islands in bladder cancer." CANCER RESEARCH, vol. 60, no. 9, 1 May 2000 (2000-05-01), pages 2473-2476, XP002227131 ISSN: 0008-5472 the whole document ---	1-12
X	LI LONG-CHENG ET AL: "Frequent methylation of estrogen receptor in prostate cancer: Correlation with tumor progression." CANCER RESEARCH, vol. 60, no. 3, 1 February 2000 (2000-02-01), pages 702-706, XP001121059 ISSN: 0008-5472 the whole document ---	1-12
X	VERKAJK NICOLE S ET AL: "Silencing of CD44 expression in prostate cancer by hypermethylation of the CD44 promoter region." LABORATORY INVESTIGATION, vol. 80, no. 8, August 2000 (2000-08), pages 1291-1298, XP001121060 ISSN: 0023-6837 the whole document ---	1-12
X	GONZALGO M L AND JONES P A: "Rapid quantification of methylation differences at specific sites using methylation-sensitive single nucleotide primer extension (Ms-SNuPE)" NUCLEIC ACIDS RESEARCH, OXFORD UNIVERSITY PRESS, SURREY, GB, vol. 25, no. 12, 1997, pages 2529-2531, XP002106409 ISSN: 0305-1048 cited in the application the whole document ---	1-12

	-/--	

INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 01/51652

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	TOYOTA MINORU ET AL: "Identification of differentially methylated sequences in colorectal cancer by methylated CpG island amplification" CANCER RESEARCH, AMERICAN ASSOCIATION FOR CANCER RESEARCH, BALTIMORE, MD, US, vol. 59, no. 10, 15 May 1999 (1999-05-15), pages 2307-2312, XP002211911 ISSN: 0008-5472 cited in the application the whole document	1-12
A	--- SZYF M: "The DNA methylation machinery as a therapeutic target" CURRENT DRUG TARGETS, July 2000 (2000-07), pages 101-101-118, XP001122812 the whole document -----	1-12

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 01/51652

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1, 2, 4-12 (partially)

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

Invention 1: claims 1, 2, 4-12 (partially)

a diagnostic or prognostic assay for cancer, comprising obtaining a tissue sample from a test tissue, performing a methylation assay on DNA derived from the sample, wherein the methylation assay determines the methylation state of a CpG dinucleotide within a DNA sequence of the DNA, and wherein the DNA sequence is a sequence selected from the group consisting of SEQ ID NO 1, sequences having a nucleotide sequence at least 90% identical to sequence SEQ ID NO 1, CpG island sequences associated with SEQ ID NO 1, and combinations thereof...within the DNA sequence; a kit useful for the detection of a methylated CpG-containing nucleic acid

Inventions 2-103: claims 1, 2,
4-12 (partially) and 3 (partially and when applicable)

a diagnostic or prognostic assay for cancer, comprising obtaining a tissue sample from a test tissue, performing a methylation assay on DNA derived from the sample, wherein the methylation assay determines the methylation state of a CpG dinucleotide within a DNA sequence of the DNA, and wherein the DNA sequence is a sequence selected from the group consisting of SEQ ID NO 2-103, sequences having a nucleotide sequence at least 90% identical to sequence SEQ ID NO 1, CpG island sequences associated with SEQ ID NO 2-103, and combinations thereof...within the DNA sequence; a kit useful for the detection of a methylated CpG-containing nucleic acid

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 01/51652

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
WO 0001816	A	13-01-2000	WO 0001816 A1	13-01-2000
			AU 8229598 A	24-01-2000

US 5552277	A	03-09-1996	AT 226643 T	15-11-2002
			AU 695420 B2	13-08-1998
			AU 3134195 A	16-02-1996
			CA 2195396 A1	01-02-1996
			DE 69528653 D1	28-11-2002
			DK 771362 T3	17-02-2003
			EP 0771362 A1	07-05-1997
			JP 10504187 T	28-04-1998
			WO 9602674 A1	01-02-1996
